Extension preservation in the finite and prefix classes of first order logic

Abhisekh Sankaran
University of Cambridge

Joint work with Anuj Dawar

YR-OWLS
Nov 11, 2020

Introduction

- Extension preservation is a well realized property in computer science. E.g. graphs containing a triangle, of chromatic number ≥ 6, of clique-width ≥ 10, etc.
- The Łoś-Tarski theorem (1954 – 55) characterizes FO definable extension preserved properties of arbitrary structures in terms of existential sentences.
- Historically significant: among the earliest applications of Gödel's Compactness theorem and opened the area of preservation theorems in model theory.
- Fails in the finite: there is an extension preserved FO sentence that is not equivalent to any existential sentence over all finite structures (Tait, 1959).

Main results

Let $\Sigma_n := \underbrace{\exists \bar{x}_1 \forall \bar{x}_2 \exists \bar{x}_3 \ldots}_{n \text{ blocks}} \alpha(\bar{x}_1, \ldots, \bar{x}_n)$ where α is quantifier-free.

Theorem

Tait's counterexample is a Σ_3 FO sentence that is extension preserved over all finite structures, but is not equivalent over this class to any Π_3 sentence. Further, the counterexample can be expressed in Datalog$(\neq, \neg)$.

Theorem

For every n, there is a vocabulary σ_n and an FO(σ_n) Σ_{2n+1} sentence φ_n that is extension closed over all finite structures, but that is not equivalent over this class to any Π_{2n+1} sentence. Further, φ_n can be expressed in Datalog$(\neq, \neg)$.

Main results

Let $\Sigma_n := \underbrace{\exists \bar{x}_1 \forall \bar{x}_2 \exists \bar{x}_3 \ldots}_{n \text{ blocks}} \alpha(\bar{x}_1, \ldots, \bar{x}_n)$ where α is quantifier-free.

Theorem

Tait's counterexample is a Σ_3 FO sentence that is extension preserved over all finite structures, but is not equivalent over this class to any Π_3 sentence. Further, the counterexample can be expressed in Datalog($\neq, \neg$).

Theorem

No prefix class of FO is expressive enough to capture:

- Extension closed FO properties in the finite
- FO $\cap$ Datalog($\neq, \neg$) queries in the finite

Part I: Analysing Tait's sentence

Overview

- The sentence SomeTotalR
- Datalog($\neq, \neg$) definition
- Non-preservation under extensions in the infinite
- Inexpressibility in Π_3 via construction of a suitable model and non-model of SomeTotalR

The sentence

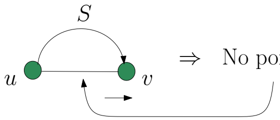
Tait's sentence

$$\text{SomeTotalR} := (\text{LO} \wedge \text{PartialSucc}) \rightarrow \exists u \exists v \ \text{RTotal}(u, v)$$

$(\in \text{FO}(\sigma)$ where $\sigma = \{\leq, R, S\})$

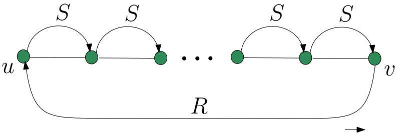
LO := "$\leq$ is a linear order"

PartialSucc := $\forall u \forall v$



$\Rightarrow$ No points between u and v

RTotal(u, v) :=



Tait's sentence

$$\mathrm{SomeTotalR} := (\mathrm{LO} \wedge \mathrm{PartialSucc}) \to \exists u \exists v\ \mathrm{RTotal}(u, v)$$

$$(\in \mathrm{FO}(\sigma) \text{ where } \sigma = \{\leq, R, S\})$$

$$\mathrm{LO} := \quad \forall x \forall y \forall z \quad \begin{pmatrix} x \leq x \ \wedge \\ (x \leq y \ \wedge \ y \leq x) \to \ x = y \ \wedge \\ (x \leq y \ \wedge \ y \leq z) \to \ x \leq z \end{pmatrix}$$

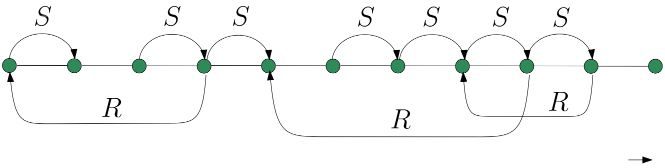
$$\mathrm{PartialSucc} :=$$

$$\forall u \forall v S(u, v) \to \forall z (z \leq u \vee v \leq z)$$

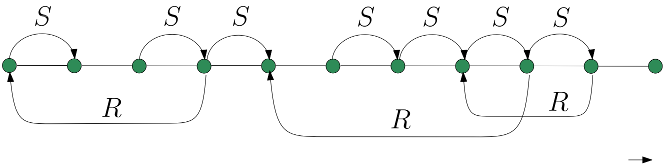
$$\mathrm{RTotal}(u, v) :=$$

$$R(v, u) \wedge \Big(\forall z (u \leq z \wedge z < v) \to \exists w (z < w \wedge w \leq v \wedge S(z, w)) \Big)$$

A model for SomeTotalR

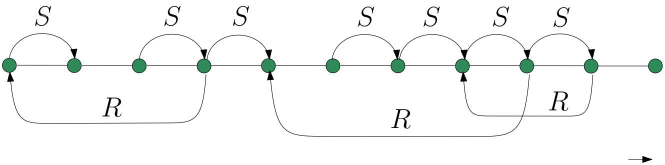


A model for SomeTotalR



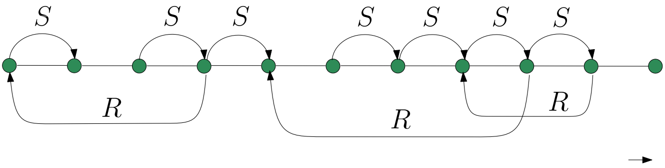
LO:= "$\leq$ is a linear order"

A model for SomeTotalR

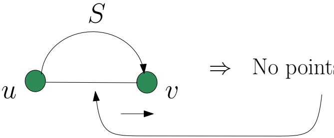


LO:= "$\leq$ is a linear order" ✓

A model for SomeTotalR

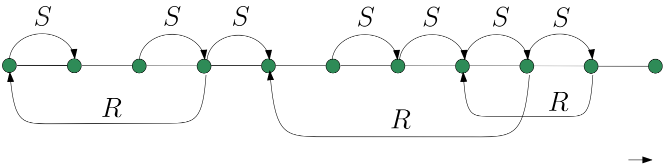


PartialSucc $:= \forall u \forall v$

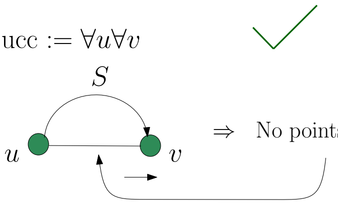


$\Rightarrow$ No points between u and v

A model for SomeTotalR

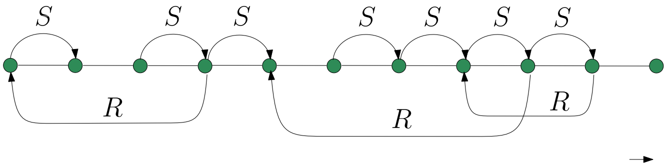


PartialSucc := $\forall u \forall v$

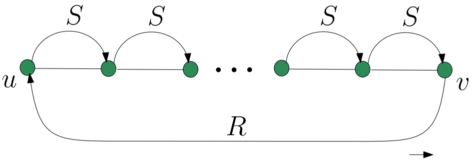


$\Rightarrow$ No points between u and v

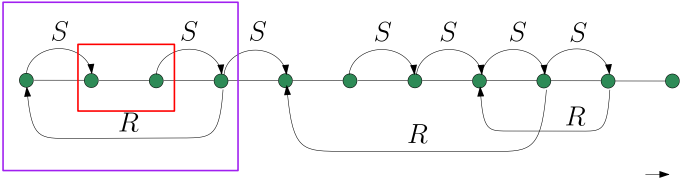
A model for SomeTotalR



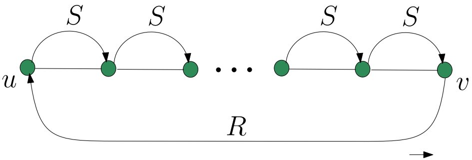
RTotal(u, v) :=



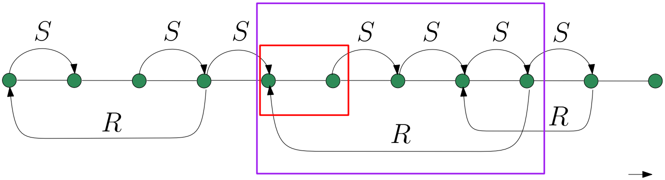
A model for SomeTotalR



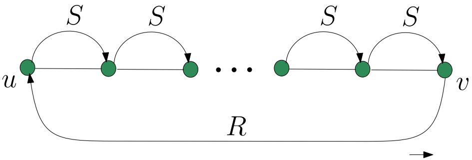
$\mathrm{RTotal}(u, v) :=$



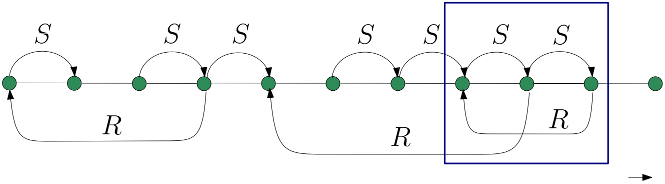
A model for SomeTotalR



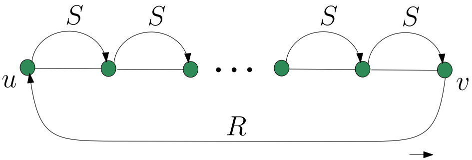
$\mathrm{RTotal}(u, v) :=$



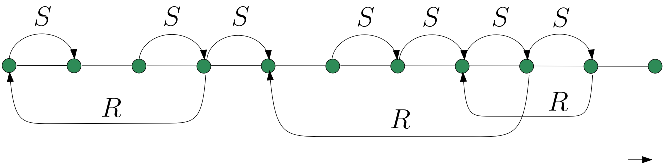
A model for SomeTotalR



$\mathrm{RTotal}(u, v) :=$



A model for SomeTotalR



$$\text{SomeTotal} := (\text{LO} \wedge \text{PartialSucc}) \rightarrow \exists u \exists v \ \text{RTotal}(u, v)$$

✓

Datalog($\neq, \neg$) definition of Tait's sentence

Datalog($\neq, \neg$) syntax

- A Datalog($\neq, \neg$) rule is of one of the foll. forms:

$$\begin{array}{lcl} R(\bar{x}) & \longleftarrow & A(\bar{x}_1) \\ R(\bar{x}) & \longleftarrow & R_1(\bar{x}_1), \ldots, R_n(\bar{x}_n) \end{array}$$

- In the first rule above, $A(\bar{x}_1)$ is an atom that can appear negated. Also A can be equality or its negation.
- In the second rule above, all predicates R_i that are not atoms appear un-negated. Also, R can be one of the R_is.
- In both rules, the variables appearing in the LHS are a subset of the variables appearing in the RHS.
- A Datalog($\neq, \neg$) program is a finite set of Datalog($\neq, \neg$) rules.

Datalog($\neq, \neg$) semantics

- Consider the following Datalog($\neq, \neg$) program:

$$\begin{array}{rcl} R(x,y) & \longleftarrow & A(x,z), B(z,y) \\ R(x,y) & \longleftarrow & \neg A(x,z), R(x,y) \end{array}$$

- The first rule as a program by itself corresponds to

$$\alpha(x,y) := \exists z (A(x,z) \wedge B(z,y))$$

- With both rules, the program corresponds to the existential least fixpoint logic sentence $\beta(x,y)$ given as below:

$$\begin{array}{rcl} \beta(x,y) & := & \mathsf{LFP}_{R,u,v}\varphi(R,u,v)](x,y) \\ \varphi(R,u,v) & := & \alpha(u,v) \vee \exists z(\neg A(u,z) \wedge R(u,v)) \end{array}$$

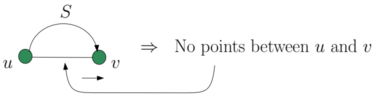
- Datalog($\neq, \neg$) corresponds exactly to existential least fixpoint logic, and thus any Datalog($\neq, \neg$) program is extension closed.

SomeTotalR as a Datalog($\neq, \neg$) program

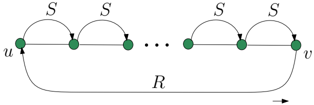
$$\text{SomeTotalR} := (\text{LO} \wedge \text{PartialSucc}) \rightarrow \exists u \exists v \ \text{RTotal}(u, v)$$

LO := "$\leq$ is a linear order"

PartialSucc := $\forall u \forall v$



RTotal(u, v) :=



- Express $\neg$LO, $\neg$PartialSucc, $\exists u \exists v$RTotal(u, v) as Datalog($\neq, \neg$) programs with "start symbols" NotLO, NotPartialSucc, RTotal(u, v) resp. Then the Datalog($\neq, \neg$) program for SomeTotalR is

$$\text{SomeTotalR} \longleftarrow \text{NotLO} \mid \text{NotPartialSucc} \mid \text{RTotal}(u, v)$$

SomeTotalR as a Datalog($\neq, \neg$) program

LO := "$\leq$ is a linear order"

$$\text{LO} := \forall x \forall y \forall z \quad \begin{pmatrix} x \leq x \quad \wedge \\ (x \leq y \;\wedge\; y \leq x) \rightarrow \; x = y \;\wedge \\ (x \leq y \;\wedge\; y \leq z) \rightarrow \; x \leq z \end{pmatrix}$$

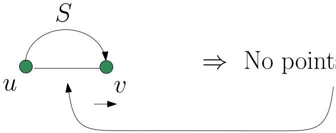
$$\neg\, \text{LO} := \exists x \exists y \exists z \quad \begin{pmatrix} \neg\, x \leq x \qquad \vee \\ (x \leq y \wedge y \leq x \wedge x \neq y) \quad \vee \\ (x \leq y \wedge y \leq z \wedge \neg\, x \leq z) \end{pmatrix}$$

Datalog($\neq, \neg$) program for $\neg$ LO:

$$\begin{aligned} \text{NotLO} \longleftarrow\; & \neg x \leq x \mid \\ & x \leq y,\; y \leq x,\; x \neq y \mid \\ & x \leq y,\; y \leq z,\; \neg\, x \leq z \end{aligned}$$

SomeTotalR as a Datalog($\neq, \neg$) program

PartialSucc := $\forall u \forall v$



$\Rightarrow$ No points in here

$$\text{PartialSucc} := \forall u \forall v\ S(u,v) \rightarrow \neg \exists z \begin{pmatrix} u \leq z \wedge z \leq v \ \wedge \\ u \neq z \wedge z \neq v \end{pmatrix}$$

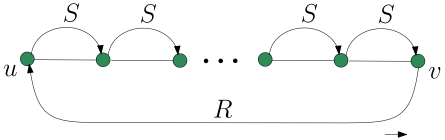
$$\neg\text{PartialSucc} := \exists u \exists v\ S(u,v) \wedge \exists z \begin{pmatrix} u \leq z \wedge z \leq v \ \wedge \\ u \neq z \wedge z \neq v \end{pmatrix}$$

Datalog($\neq, \neg$) program for $\neg$PartialSucc:

$$\text{NotPartialSucc} \longleftarrow S(u,v), X(u,v)$$
$$X(u,v) \longleftarrow u \leq z, z \leq v, u \neq z, z \neq v$$

SomeTotalR as a Datalog($\neq, \neg$) program

$\mathrm{RTotal}(u, v) :=$



$$\mathrm{RTotal}(u, v) := R(v, u) \wedge \forall z (u \leq z \wedge z < v) \rightarrow \exists w (z < w \wedge w \leq v \wedge S(z, w))$$

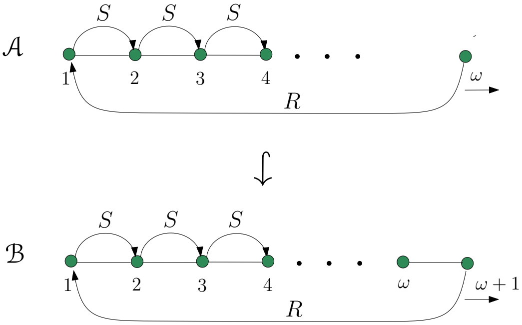
Datalog($\neq, \neg$) program for $\mathrm{RTotal}(u, v)$:

$$\mathrm{RTotal}(u, v) \longleftarrow R(v, u), S\text{-reach}(u, v)$$

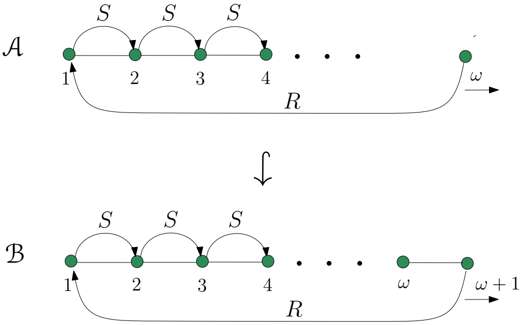
$$S\text{-reach}(u, v) \longleftarrow S(u, v) \;\mid\; S(u, z), S\text{-reach}(z, v)$$

Non extension preservation of Tait's sentence in the infinite

SomeTotalR is not extension preserved in the infinite

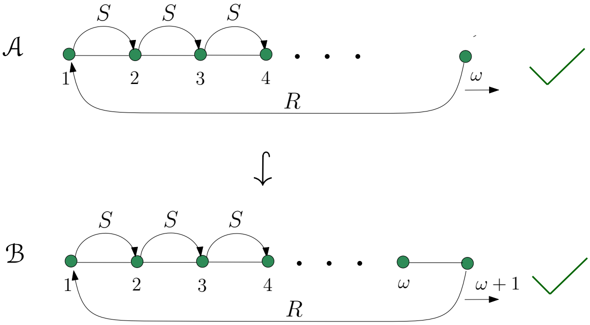


SomeTotalR is not extension preserved in the infinite



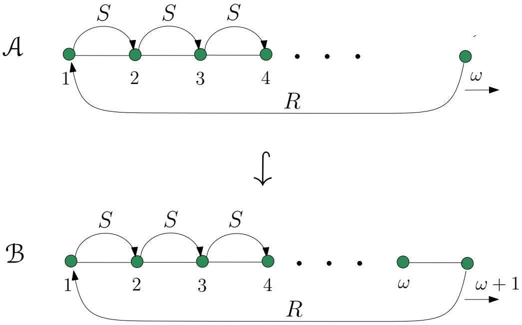
LO := "$\leq$ is a linear order"

SomeTotalR is not extension preserved in the infinite



LO := "$\leq$ is a linear order"

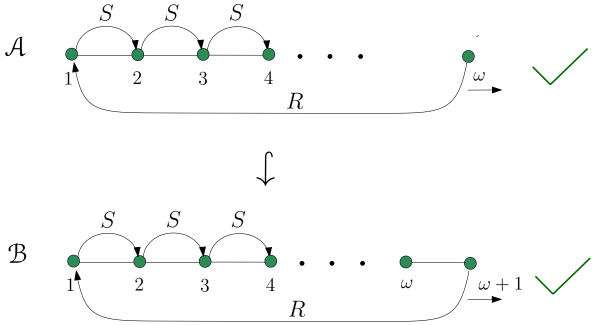
SomeTotalR is not extension preserved in the infinite



PartialSucc :=

$$\forall u \forall v S(u, v) \rightarrow \forall z (z \leq u \vee v \leq z)$$

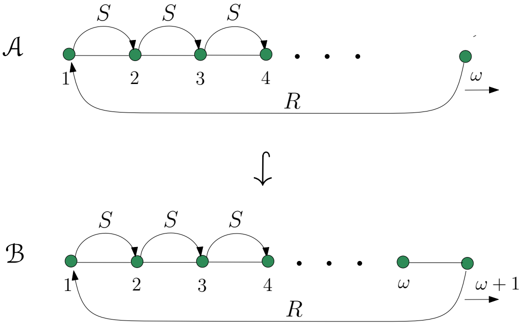
SomeTotalR is not extension preserved in the infinite



PartialSucc :=

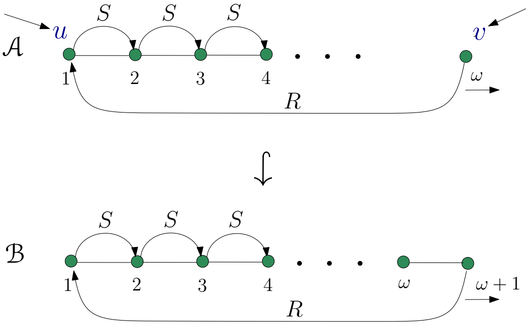
$$\forall u \forall v S(u, v) \rightarrow \forall z (z \leq u \vee v \leq z)$$

SomeTotalR is not extension preserved in the infinite



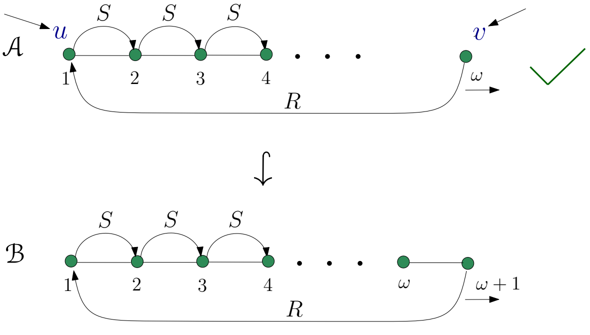
$$\text{RTotal}(u, v) := R(v, u) \wedge \Big(\forall z(u \leq z \wedge z < v) \rightarrow \exists w(z < w \wedge w \leq v \wedge S(z, w))\Big)$$

SomeTotalR is not extension preserved in the infinite



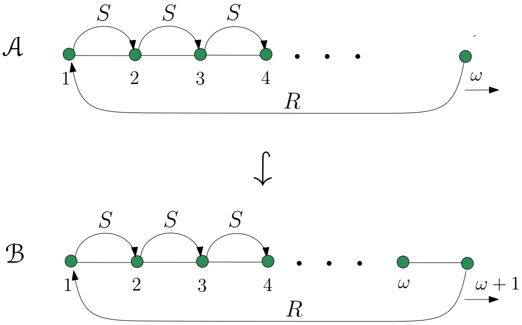
$$\mathrm{RTotal}(u, v) :=$$
$$R(v, u) \wedge \Big(\forall z(u \leq z \wedge \boxed{z < v}) \rightarrow \exists w(z < w \wedge w \leq v \wedge S(z, w))\Big)$$

SomeTotalR is not extension preserved in the infinite



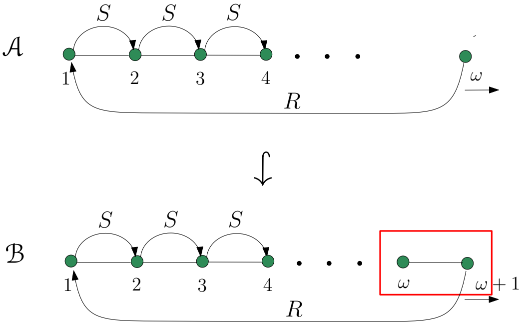
$$\mathrm{RTotal}(u, v) :=$$
$$R(v, u) \wedge \Big(\forall z(u \leq z \wedge \boxed{z < v}) \rightarrow \exists w(z < w \wedge w \leq v \wedge S(z, w))\Big)$$

SomeTotalR is not extension preserved in the infinite



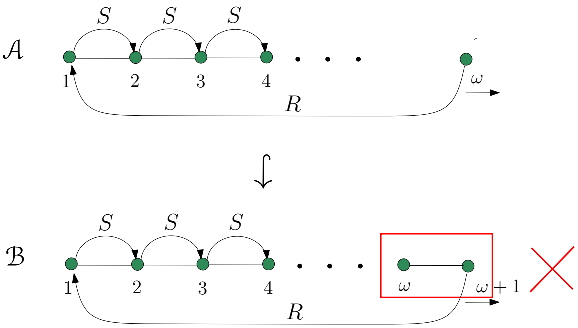
$$\mathrm{RTotal}(u, v) :=$$
$$R(v, u) \wedge \Big(\forall z (u \leq z \wedge z < v) \rightarrow \exists w (z < w \wedge w \leq v \wedge S(z, w))\Big)$$

SomeTotalR is not extension preserved in the infinite



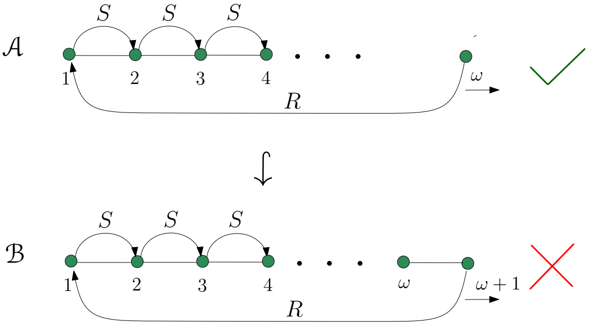
$$\mathrm{RTotal}(u, v) :=$$
$$R(v, u) \wedge \Big(\forall z(u \leq z \wedge z < v) \rightarrow \exists w(z < w \wedge w \leq v \wedge S(z, w))\Big)$$

SomeTotalR is not extension preserved in the infinite



$$\mathrm{RTotal}(u, v) := R(v, u) \wedge \Big(\forall z (u \leq z \wedge z < v) \rightarrow \exists w (z < w \wedge w \leq v \wedge S(z, w))\Big)$$

SomeTotalR is not extension preserved in the infinite



$$\text{SomeTotalR} := (\text{LO} \wedge \text{PartialSucc}) \rightarrow \exists u \exists v\ \text{RTotal}(u, v)$$

Stronger failure of Łoś-Tarski theorem in the finite

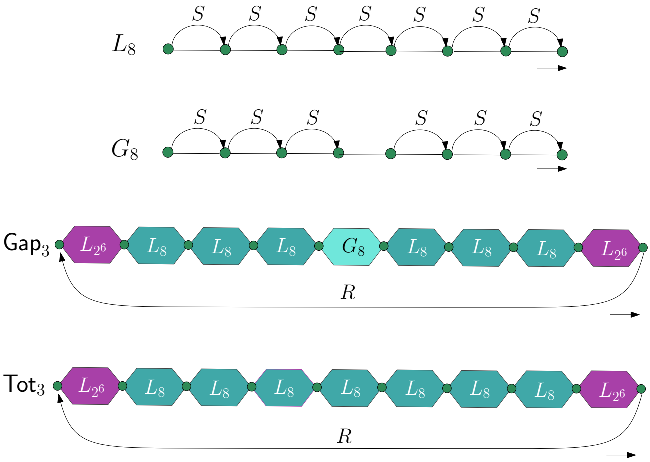
Theorem

The Σ_3 sentence SomeTotalR is not equivalent over all finite σ-structures to any Π_3 sentence.

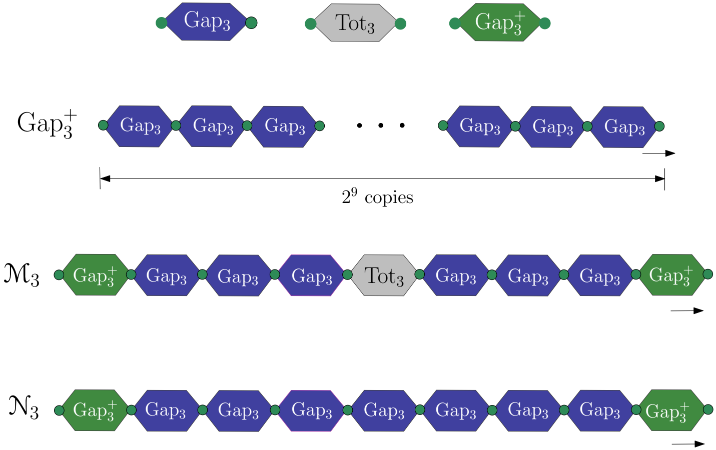
- Let $\Pi_{n,k}$ = class of all Π_n sentences in which each block of quantifiers has size k. So $\Pi_n = \bigcup_{k \geq 0} \Pi_{n,k}$.
- Let $\mathcal{A} \Rrightarrow_{n,k} \mathcal{B}$ = for each $\Sigma_{n,k}$ sentence θ, it holds that $\mathcal{A} \models \theta \rightarrow \mathcal{B} \models \theta$.
- $\mathcal{A} \Rrightarrow_{n,k} \mathcal{B}$ is equivalent to: for each $\Pi_{n,k}$ sentence γ, it holds that $\mathcal{B} \models \gamma \rightarrow \mathcal{A} \models \gamma$.
- For each k, we construct a model $\mathcal{M}_k$ and a non-model $\mathcal{N}_k$ of SomeTotalR such that $\mathcal{N}_k \Rrightarrow_{3,k} \mathcal{M}_k$.
- We illustrate our constructions for $k = 3$.

Construction of $\mathcal{M}_k$ and $\mathcal{N}_k$

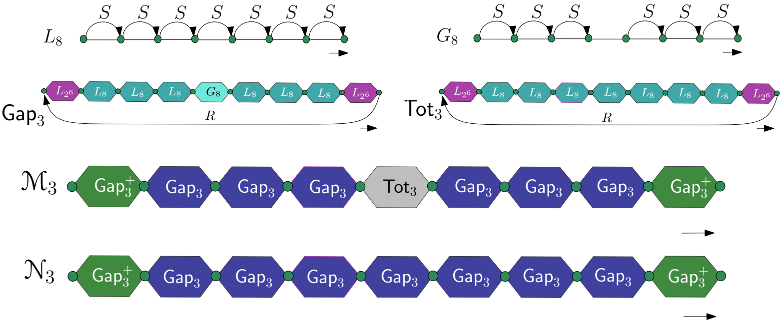
Construction of $\mathcal{M}_3$ and $\mathcal{N}_3$



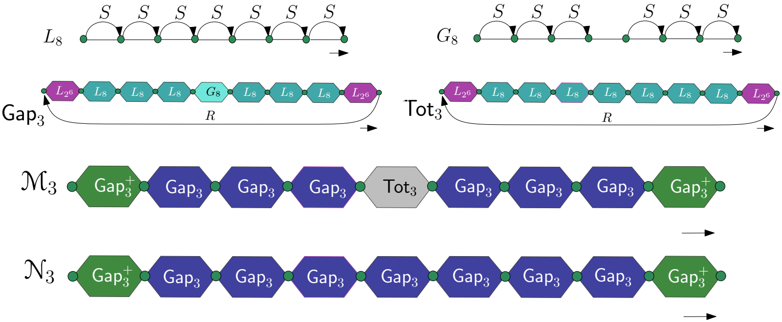
Construction of $\mathcal{M}_3$ and $\mathcal{N}_3$



$\mathcal{M}_3$ models SomeTotalR but $\mathcal{N}_3$ does not

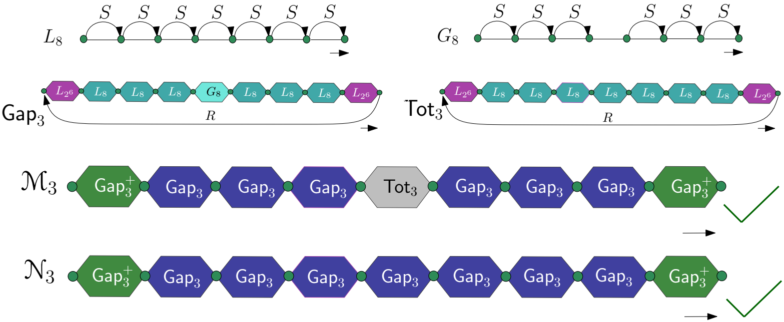


$\mathcal{M}_3$ models SomeTotalR but $\mathcal{N}_3$ does not



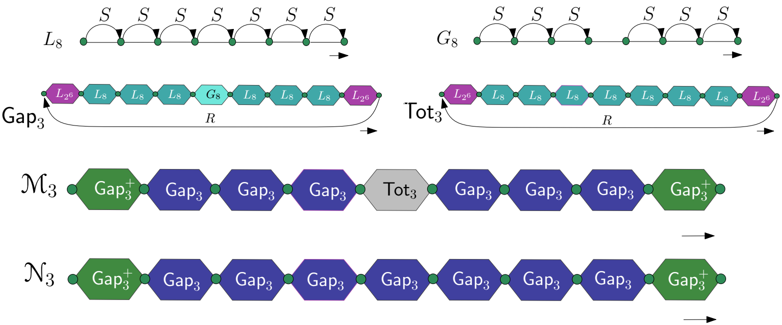
LO := "$\leq$ is a linear order"

$\mathcal{M}_3$ models SomeTotalR but $\mathcal{N}_3$ does not

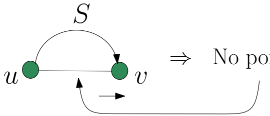


LO := "$\leq$ is a linear order"

$\mathcal{M}_3$ models SomeTotalR but $\mathcal{N}_3$ does not

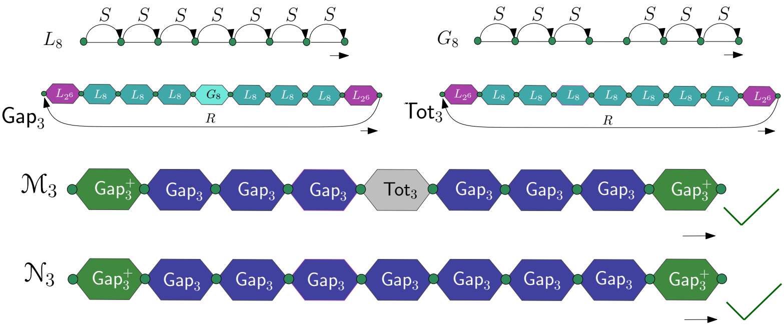


$\text{PartialSucc} := \forall u \forall v$

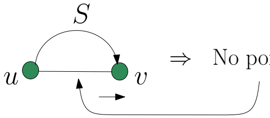


$\Rightarrow$ No points between u and v

$\mathcal{M}_3$ models SomeTotalR but $\mathcal{N}_3$ does not

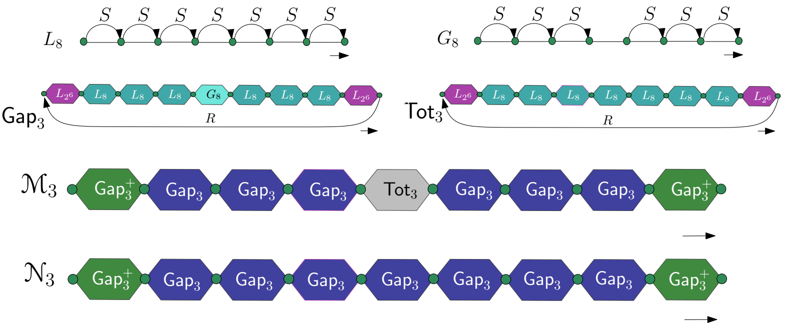


PartialSucc := $\forall u \forall v$

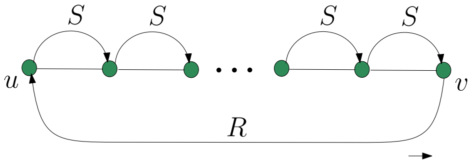


$\Rightarrow$ No points between u and v

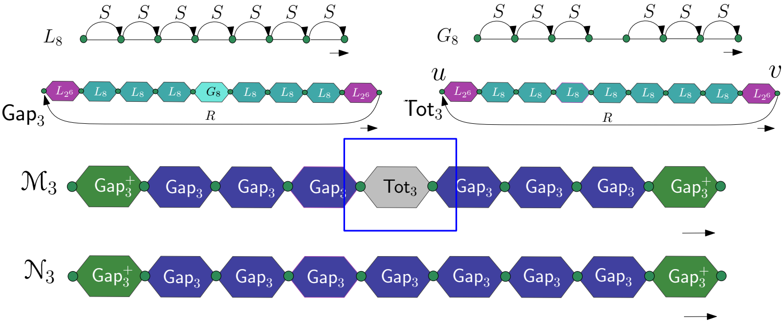
$\mathcal{M}_3$ models SomeTotalR but $\mathcal{N}_3$ does not



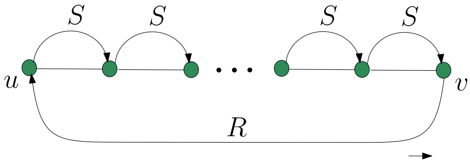
$\text{RTotal}(u, v) :=$



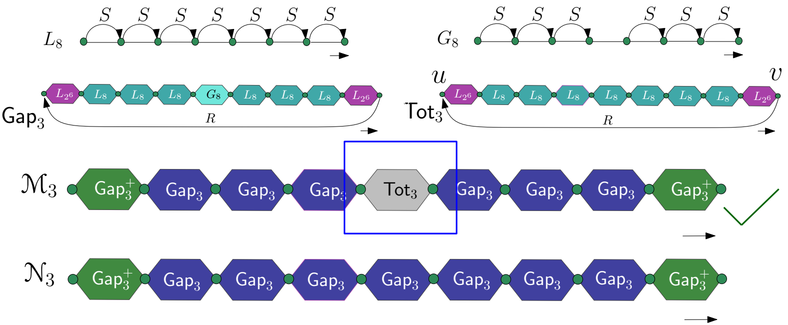
$\mathcal{M}_3$ models SomeTotalR but $\mathcal{N}_3$ does not



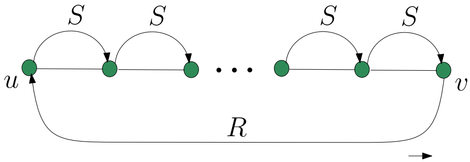
$\text{RTotal}(u, v) :=$



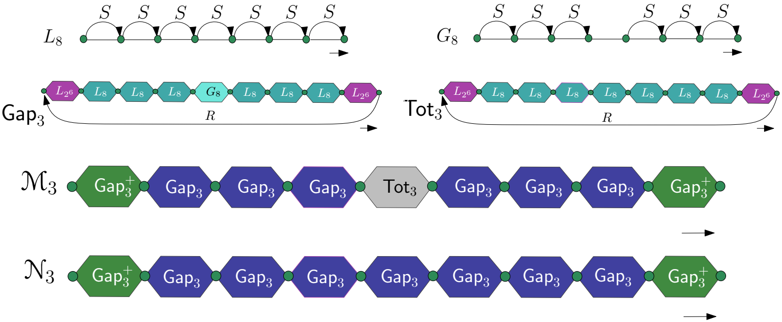
$\mathcal{M}_3$ models SomeTotalR but $\mathcal{N}_3$ does not



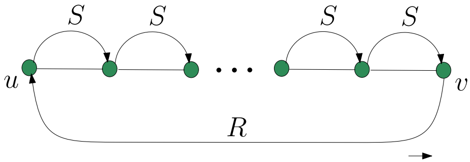
$\mathrm{RTotal}(u, v) :=$



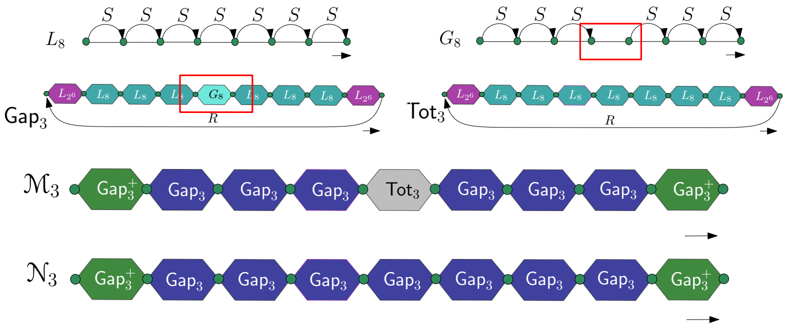
$\mathcal{M}_3$ models SomeTotalR but $\mathcal{N}_3$ does not



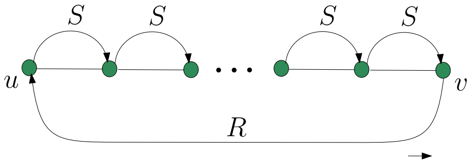
$\mathrm{RTotal}(u, v) :=$



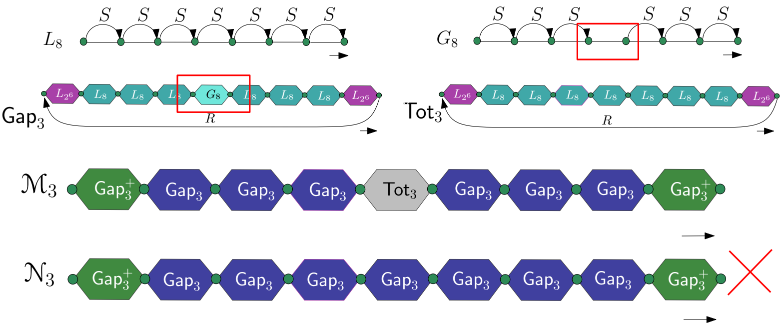
$\mathcal{M}_3$ models SomeTotalR but $\mathcal{N}_3$ does not



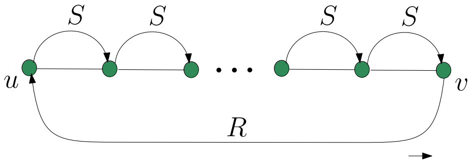
$\mathrm{RTotal}(u, v) :=$



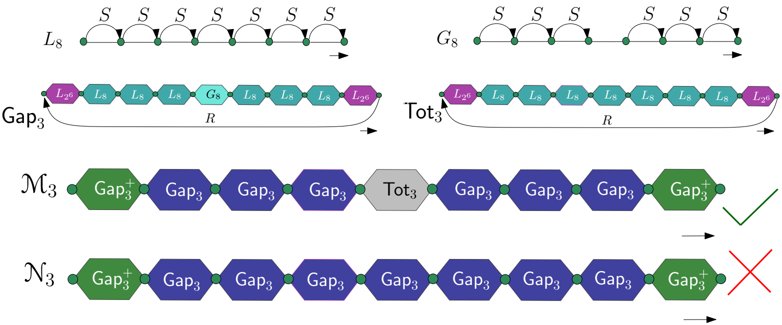
$\mathcal{M}_3$ models SomeTotalR but $\mathcal{N}_3$ does not



$\text{RTotal}(u, v) :=$



$\mathcal{M}_3$ models SomeTotalR but $\mathcal{N}_3$ does not



$$\mathrm{SomeTotalR} := (\mathrm{LO} \wedge \mathrm{PartialSucc}) \rightarrow \exists u \exists v\ \mathrm{RTotal}(u, v)$$

Inexpressibility of SomeTotalR in Π_3
via showing $\mathcal{N}_k \Rrightarrow_{3,k} \mathcal{M}_k$

Ehrenfeucht-Fraïssé (EF) game for $\Rrightarrow_{n,k}$

- Two players: Spoiler and Duplicator; Game arena: a pair $(\mathcal{A}, \mathcal{B})$ of structures; Rounds: n.
- In odd rounds i, Spoiler chooses a k-tuple $\bar{a}_i$ from $\mathcal{A}$ and in even rounds i, he chooses a k-tuple $\bar{b}_i$ from $\mathcal{B}$.
- Duplicator responds with k-tuples $\bar{b}_i$ from $\mathcal{B}$ in odd rounds and with k-tuples $\bar{a}_i$ from $\mathcal{A}$ in even rounds.
- Duplicator wins the play of the game if $(\bar{a}_i \mapsto \bar{b}_i)_{1 \leq i \leq n}$ is a partial isomorphism between $\mathcal{A}$ and $\mathcal{B}$. She has a winning strategy if she wins every play of the game.

Theorem

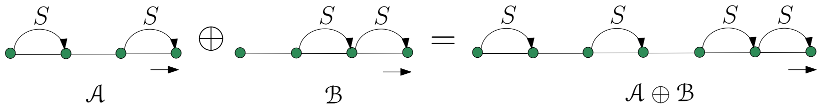
Duplicator has a winning strategy in the above game iff $\mathcal{A} \Rrightarrow_{n,k} \mathcal{B}$.

Ordered Sum

Definition

For ordered structures $\mathcal{A}$ and $\mathcal{B}$, the ordered sum $\mathcal{A} \oplus \mathcal{B}$ is the ordered structure that is the disjoint union of $\mathcal{A}$ and $\mathcal{B}$ with the additional constraints that:

- the elements of $\mathcal{A}$ appear "before" those of $\mathcal{B}$, and
- the last element of $\mathcal{A}$ is identified with the first element of $\mathcal{B}$.



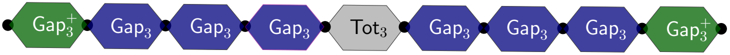
Ordered Sum

Definition

For ordered structures $\mathcal{A}$ and $\mathcal{B}$, the ordered sum $\mathcal{A} \oplus \mathcal{B}$ is the ordered structure that is the disjoint union of $\mathcal{A}$ and $\mathcal{B}$ with the additional constraints that:

- the elements of $\mathcal{A}$ appear "before" those of $\mathcal{B}$, and
- the last element of $\mathcal{A}$ is identified with the first element of $\mathcal{B}$.

$\mathcal{M}_3$



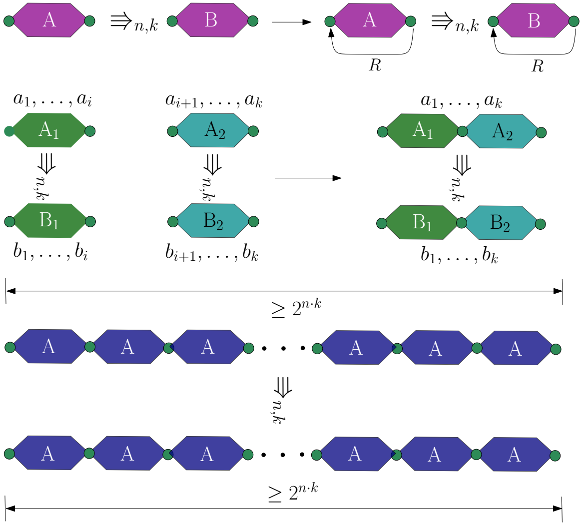
$$= \underbrace{\mathsf{Gap}_3 \oplus \ldots \oplus \mathsf{Gap}_3}_{2^9+3 \text{ copies}} \oplus \mathsf{Tot}_3 \oplus \underbrace{\mathsf{Gap}_3 \oplus \ldots \oplus \mathsf{Gap}_3}_{2^9+3 \text{ copies}}$$

Composition properties

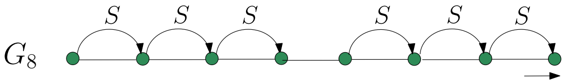
1.

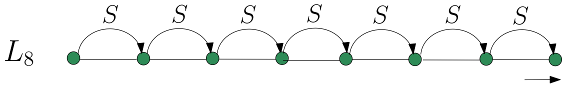
2.

3.

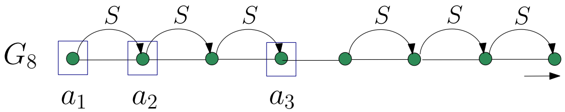


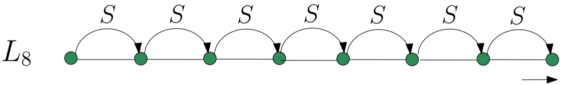
$G_8 \Rrightarrow_{1,3} L_8$



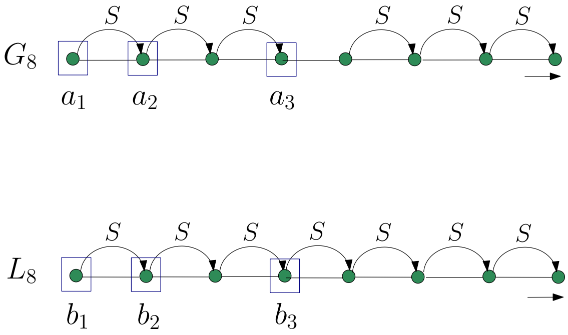


$G_8 \Rrightarrow_{1,3} L_8$

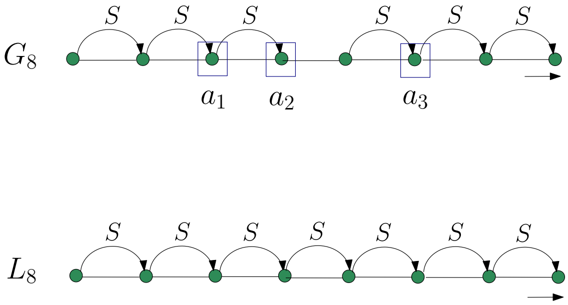




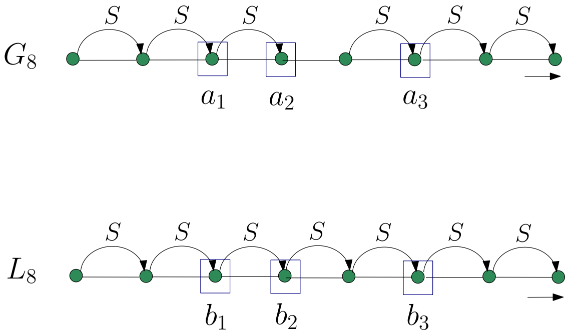
$G_8 \Rrightarrow_{1,3} L_8$



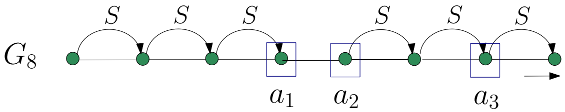
$G_8 \Rrightarrow_{1,3} L_8$

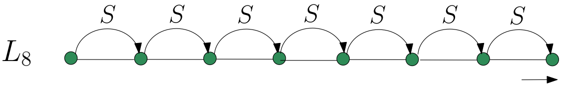


$G_8 \Rrightarrow_{1,3} L_8$

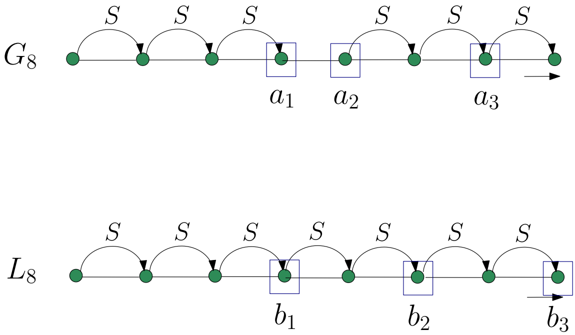


$G_8 \Rrightarrow_{1,3} L_8$

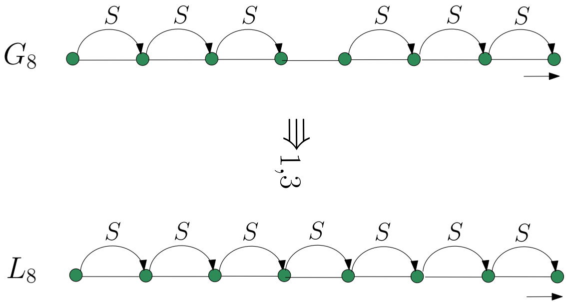




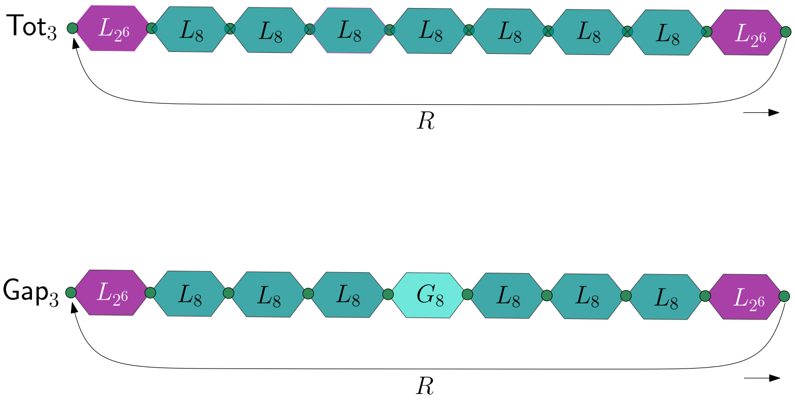
$G_8 \Rrightarrow_{1,3} L_8$



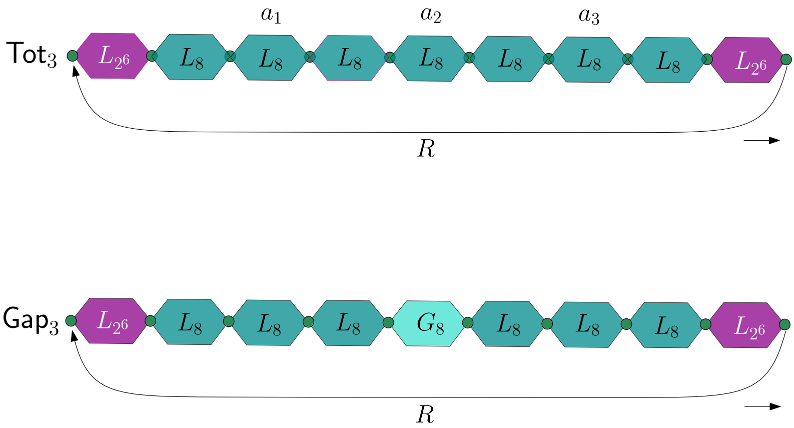
$G_8 \Rrightarrow_{1,3} L_8$



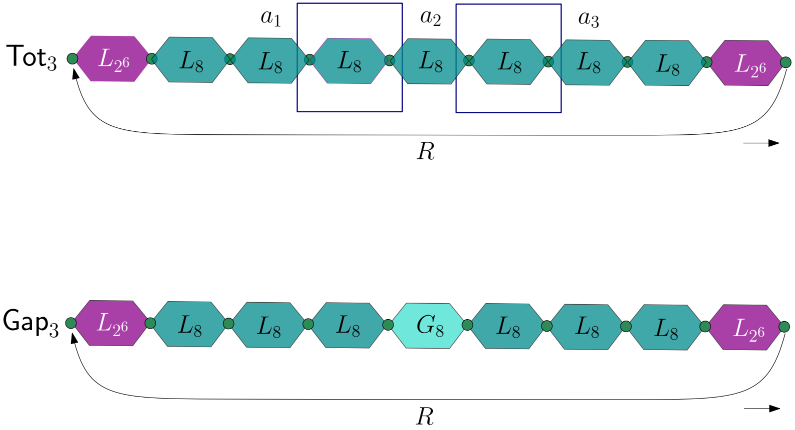
$\mathrm{Tot}_3 \Rrightarrow_{2,3} \mathrm{Gap}_3$



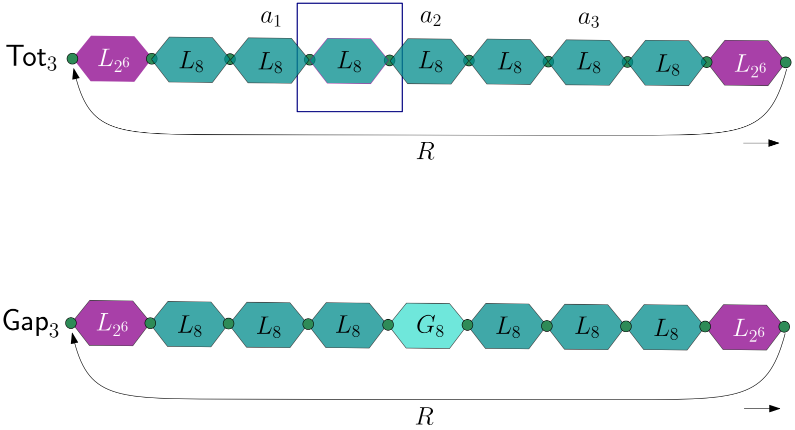
$\text{Tot}_3 \Rrightarrow_{2,3} \text{Gap}_3$



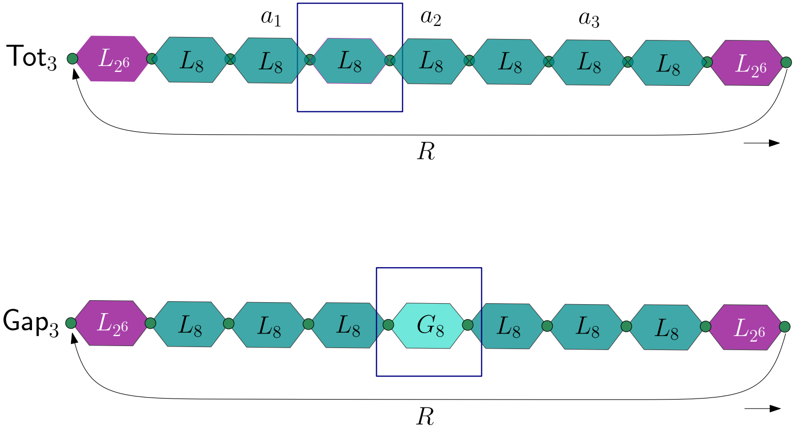
Tot$_3$ $\Rrightarrow_{2,3}$ Gap$_3$



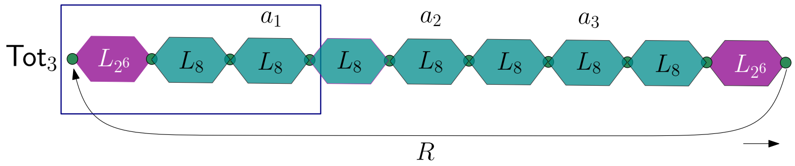
$\mathsf{Tot}_3 \Rrightarrow_{2,3} \mathsf{Gap}_3$

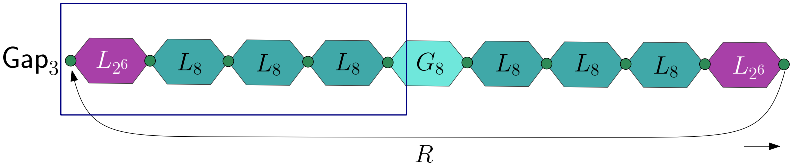


$\mathsf{Tot}_3 \Rrightarrow_{2,3} \mathsf{Gap}_3$

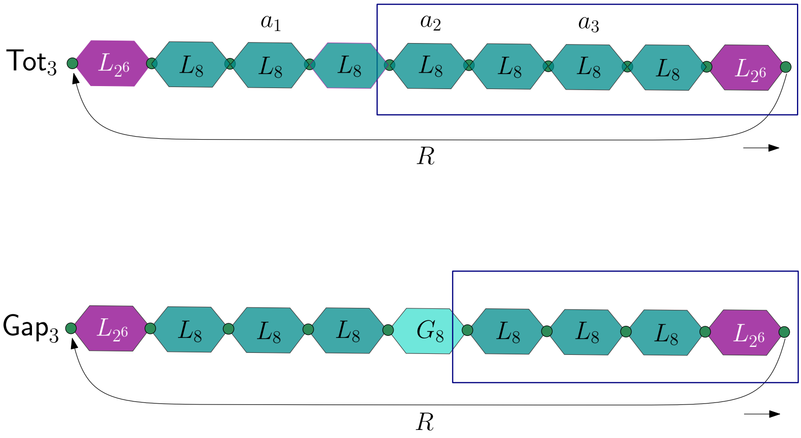


$\mathsf{Tot}_3 \Rrightarrow_{2,3} \mathsf{Gap}_3$

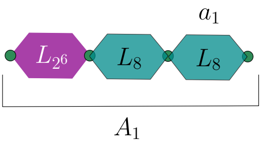


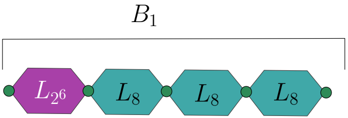


$\mathsf{Tot}_3 \Rrightarrow_{2,3} \mathsf{Gap}_3$

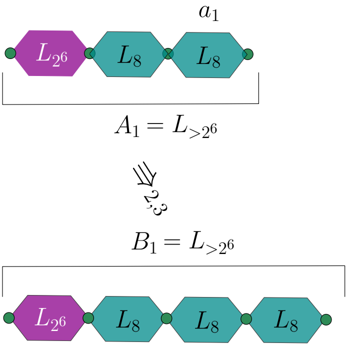


$\mathsf{Tot}_3 \Rrightarrow_{2,3} \mathsf{Gap}_3$

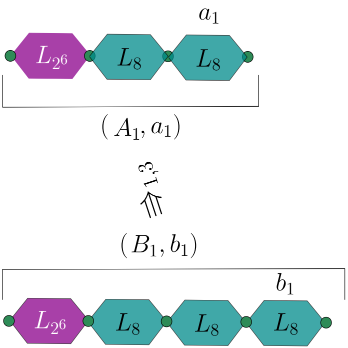




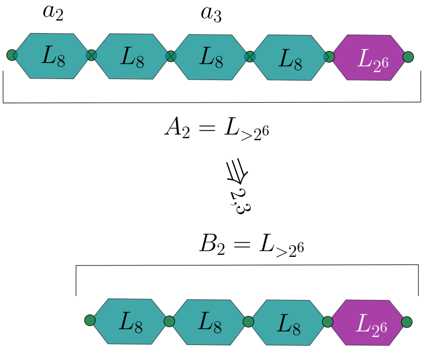
$\mathrm{Tot}_3 \Rrightarrow_{2,3} \mathrm{Gap}_3$



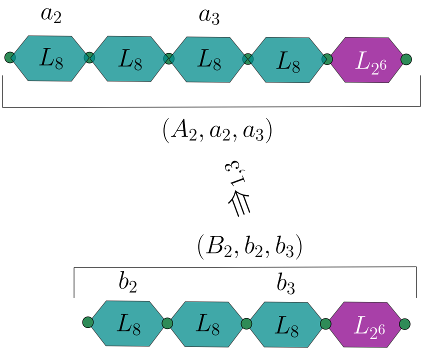
$\text{Tot}_3 \Rrightarrow_{2,3} \text{Gap}_3$



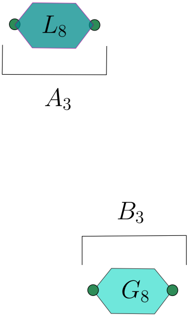
$\mathsf{Tot}_3 \Rrightarrow_{2,3} \mathsf{Gap}_3$



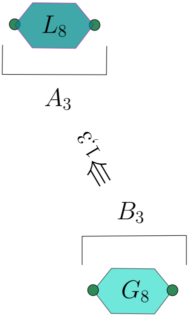
$\text{Tot}_3 \Rrightarrow_{2,3} \text{Gap}_3$



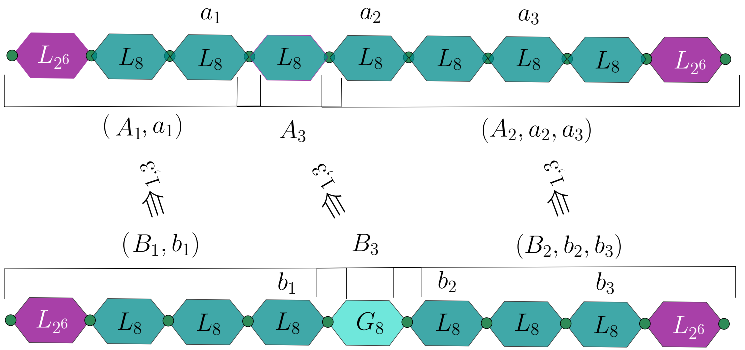
$\mathsf{Tot}_3 \Rrightarrow_{2,3} \mathsf{Gap}_3$



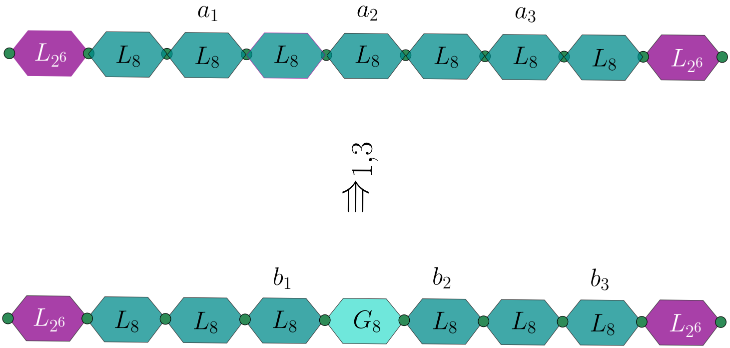
$\mathsf{Tot}_3 \Rrightarrow_{2,3} \mathsf{Gap}_3$



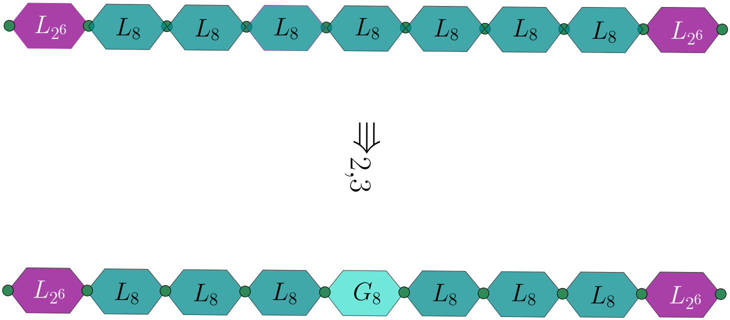
$\mathsf{Tot}_3 \Rrightarrow_{2,3} \mathsf{Gap}_3$



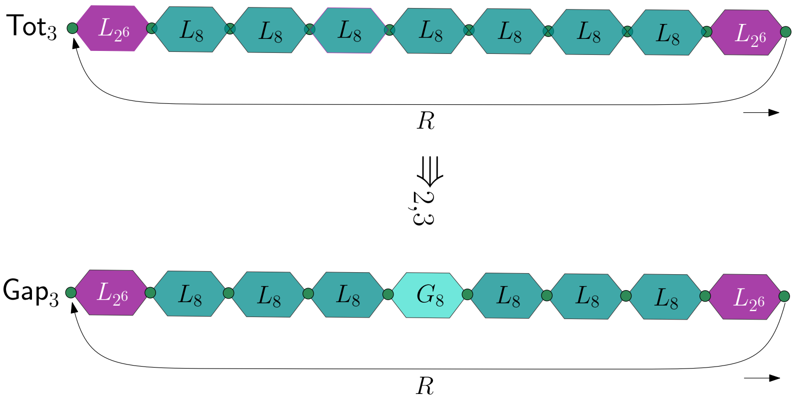
$\mathsf{Tot}_3 \Rrightarrow_{2,3} \mathsf{Gap}_3$



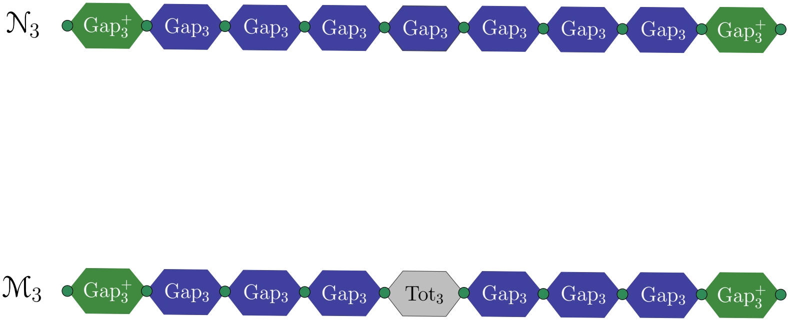
$\mathrm{Tot}_3 \Rrightarrow_{2,3} \mathrm{Gap}_3$



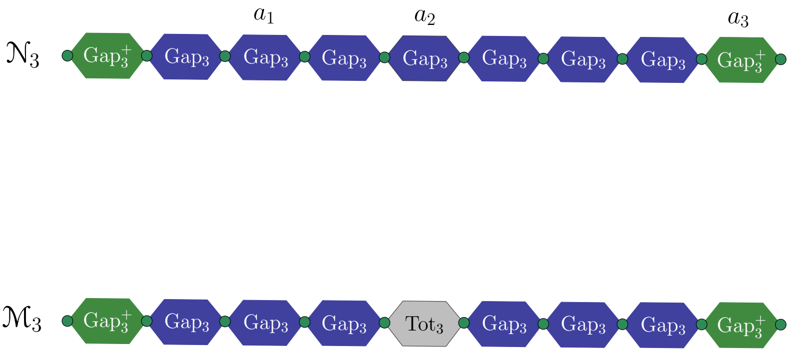
$\mathsf{Tot}_3 \Rrightarrow_{2,3} \mathsf{Gap}_3$



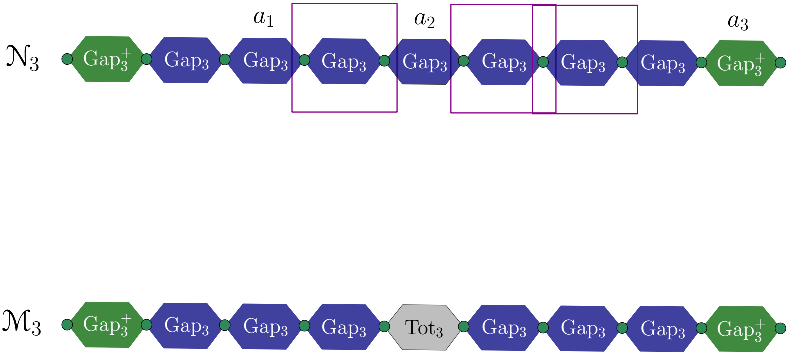
$\mathcal{N}_3 \Rrightarrow_{3,3} \mathcal{M}_3$



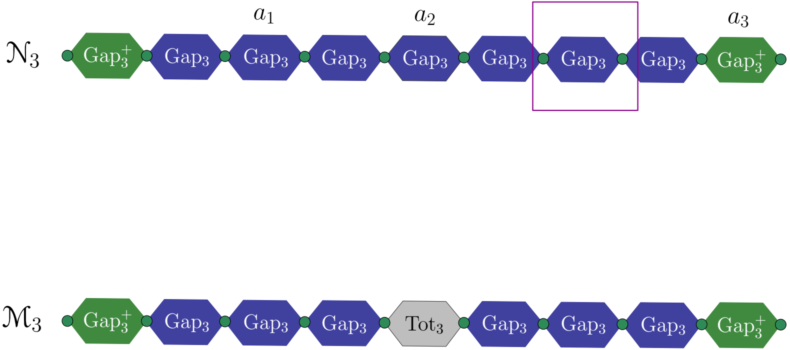
$\mathcal{N}_3 \Rrightarrow_{3,3} \mathcal{M}_3$



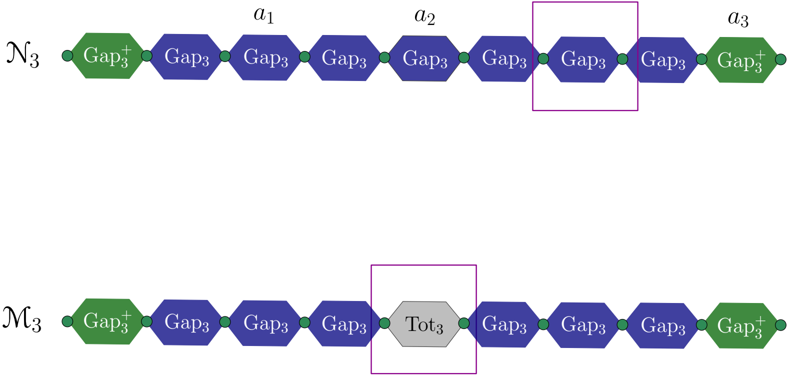
$\mathcal{N}_3 \Rrightarrow_{3,3} \mathcal{M}_3$



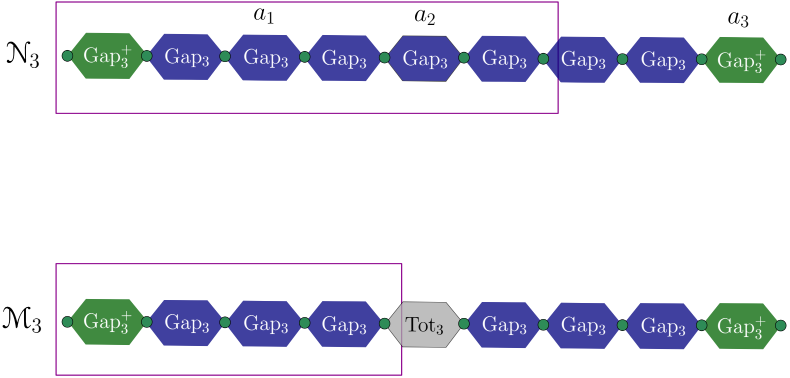
$\mathcal{N}_3 \Rrightarrow_{3,3} \mathcal{M}_3$



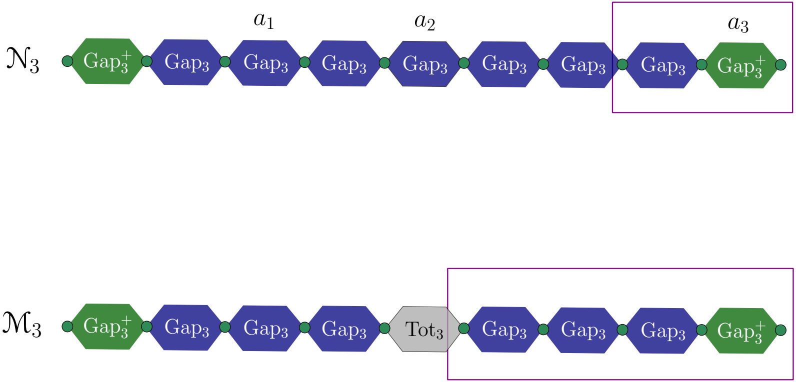
$\mathcal{N}_3 \Rrightarrow_{3,3} \mathcal{M}_3$



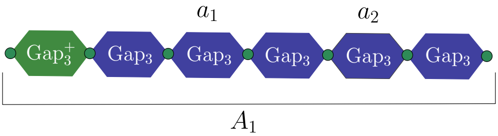
$\mathcal{N}_3 \Rrightarrow_{3,3} \mathcal{M}_3$

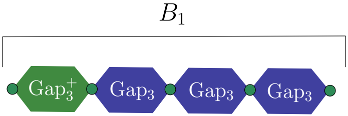


$\mathcal{N}_3 \Rrightarrow_{3,3} \mathcal{M}_3$

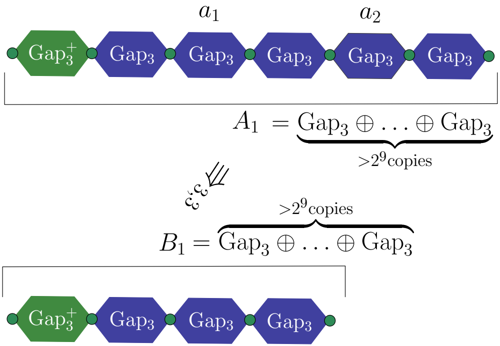


$\mathcal{N}_3 \Rrightarrow_{3,3} \mathcal{M}_3$

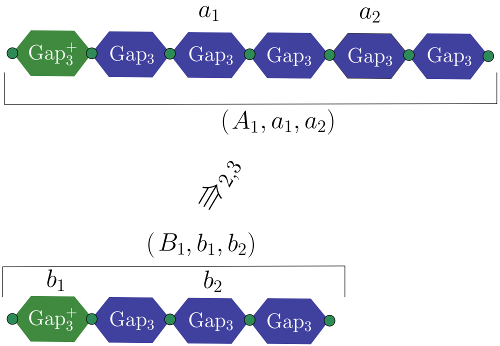




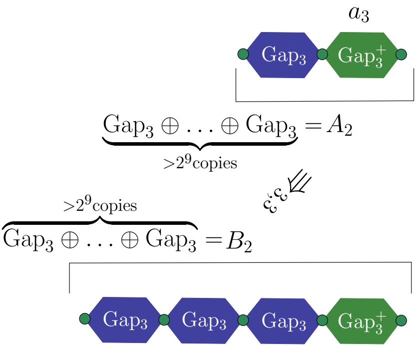
$\mathcal{N}_3 \Rrightarrow_{3,3} \mathcal{M}_3$



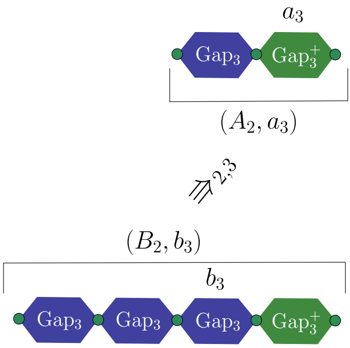
$\mathcal{N}_3 \Rrightarrow_{3,3} \mathcal{M}_3$



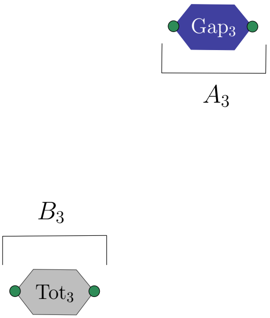
$\mathcal{N}_3 \Rrightarrow_{3,3} \mathcal{M}_3$



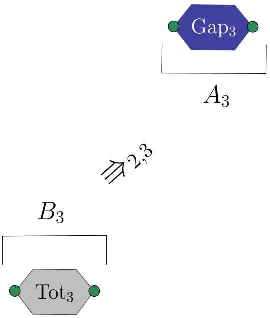
$\mathcal{N}_3 \Rrightarrow_{3,3} \mathcal{M}_3$



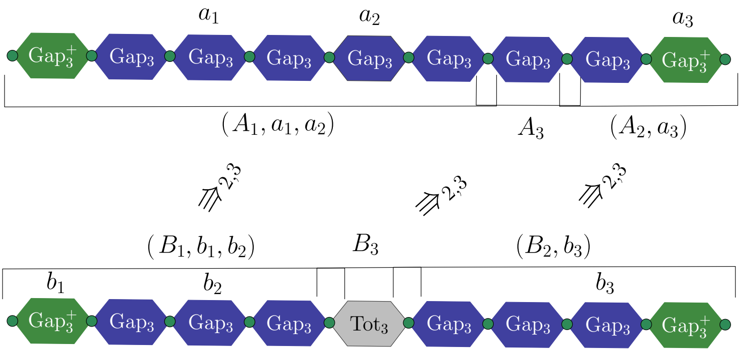
$\mathcal{N}_3 \Rrightarrow_{3,3} \mathcal{M}_3$



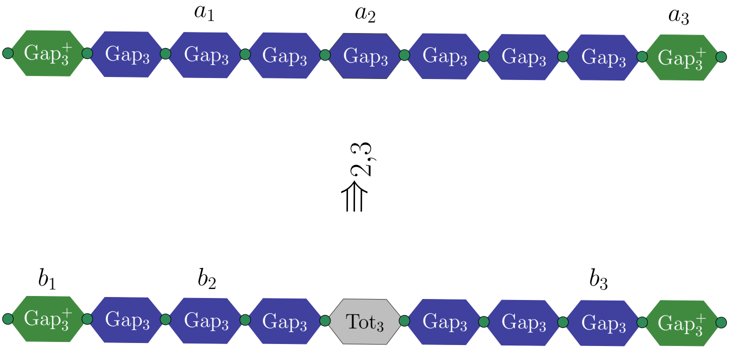
$\mathcal{N}_3 \Rrightarrow_{3,3} \mathcal{M}_3$



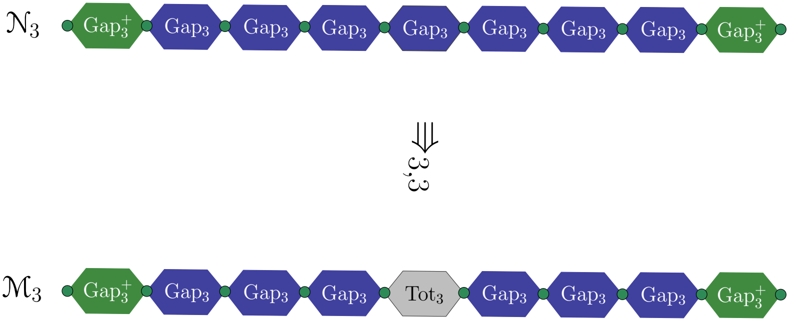
$\mathcal{N}_3 \Rrightarrow_{3,3} \mathcal{M}_3$



$\mathcal{N}_3 \Rrightarrow_{3,3} \mathcal{M}_3$



$\mathcal{N}_3 \Rrightarrow_{3,3} \mathcal{M}_3$



Generalizing Tait's sentence

Overview

Theorem

For every n, there is a vocabulary σ_n and an FO(σ_n) Σ_{2n+1} sentence SomeTotalR$_n$ such that the following hold:

1. SomeTotalR$_n$ is extension closed over all finite σ_n-structures, but is not equivalent over this class to any Π_{2n+1} sentence.
2. SomeTotalR$_n$ can be expressed in Datalog$(\neq, \neg)$.

We will see the following:

- Construction of SomeTotalR$_n$
- Datalog$(\neq, \neg)$ expressibility
- Construction of a suitable model $\mathcal{M}_{n,k}$ and non-model $\mathcal{N}_{n,k}$ such that $\mathcal{N}_{n,k} \Rrightarrow_{2n+1,k} \mathcal{M}_{n,k}$

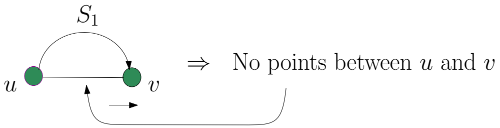
The generalized sentence SomeTotalR$_n$

Construction of SomeTotalR$_n$

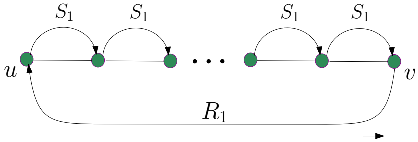
$$\text{SomeTotalR}_1 := (\text{LO} \wedge \text{PartialSucc}_1) \rightarrow \exists u \exists v\ \text{RTotal}_1(u,v)$$
$$(\in \text{FO}(\sigma_1) \text{ where } \sigma_1 = \{\leq, R_1, S_1\})$$

LO := " $\leq$ is a linear order"

PartialSucc$_1$:= $\forall u \forall v$



RTotal$_1(u,v)$:=



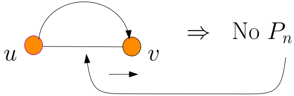
Construction of SomeTotalR$_n$

$$\text{SomeTotalR}_n := (\text{LO} \wedge \text{PartialSucc}_n) \to \exists u \exists v \ \text{RTotal}_n(u, v)$$
$$(\in \text{FO}(\sigma_n) \text{ where } \sigma_n = \sigma_{n-1} \cup \{P_n, R_n, S_n\})$$

$\text{LO} :=$ "$\leq$ is a linear order"

$\text{PartialSucc}_n := \forall u \forall v$

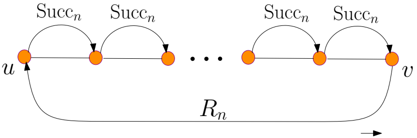
$$\text{Succ}_n := P_n(u) \wedge P_n(v) \wedge S_n(u, v) \wedge \text{SomeTotalR}_{n-1}^{[u,v]}$$



$\Rightarrow$ No P_n points between u and v

● – P_n

$\text{RTotal}_n(u, v) :=$



Construction of $\mathrm{SomeTotalR}_n$

$$\mathrm{SomeTotalR}_n := (\mathrm{LO} \wedge \mathrm{PartialSucc}_n) \rightarrow \exists u \exists v\ \mathrm{RTotal}_n(u, v)$$

$$(\in \mathrm{FO}(\sigma_n) \text{ where } \sigma_n = \sigma_{n-1} \cup \{P_n, R_n, S_n\})$$

$\mathrm{LO} :=$ " $\leq$ is a linear order"

$\mathrm{PartialSucc}_n :=$

$$\forall u \forall v \mathrm{Succ}_n(u, v) \rightarrow \forall z \big(P_n(z) \rightarrow (z \leq u \vee v \leq z)\big)$$

$\mathrm{Succ}_n(u, v) := P_n(u) \wedge P_n(v) \wedge S_n(u, v) \wedge \mathrm{SomeTotalR}_{n-1}^{[u,v]}$

$\mathrm{RTotal}_n(u, v) :=$

$$P_n(u) \wedge P_n(v) \wedge R_n(v, u) \wedge$$
$$\forall z \big((P_n(z) \wedge u \leq z \wedge z < v) \rightarrow$$
$$\exists w(P_n(w) \wedge z < w \wedge w \leq v \wedge \mathrm{Succ}_n(z, w)\big)$$

$\mathsf{SomeTotalR}_n$ as a Datalog($\neq, \neg$) program

- We construct Datalog($\neq, \neg$) programs inductively for $\mathsf{SomeTotalR}_n^{[x,y]}$ with start symbol $\mathsf{STR}_n(x, y)$.
- Then the Datalog($\neq, \neg$) program for $\mathsf{SomeTotalR}_n$ is simply

$$\mathsf{SomeTotalR}_n \longleftarrow \mathsf{STR}_n(x, y)$$

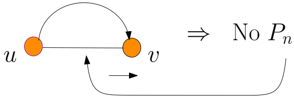
- The Datalog($\neq, \neg$) program for $\mathsf{SomeTotalR}_1$ is similar to that for Tait's sentence. (It also contains the program for $\neg$ LO).
- Assume the program for $\mathsf{SomeTotalR}_{n-1}^{[x,y]}$ has been constructed.

SomeTotalR$_n$ as a Datalog($\neq, \neg$) program

$$\text{SomeTotalR}_n := (\text{LO} \wedge \text{PartialSucc}_n) \rightarrow \exists u \exists v \ \text{RTotal}_n(u, v)$$

$\text{PartialSucc}_n := \forall u \forall v$

$\text{Succ}_n \ (:= P_n(u) \wedge P_n(v) \wedge S_n(u,v) \wedge \text{SomeTotalR}_{n-1}^{[u,v]})$



$\Rightarrow$ No P_n points between u and v

$$\neg\text{PartialSucc}_n := \exists u \exists v \ \text{Succ}_n(u, v) \wedge \exists z \begin{pmatrix} P_n(z) \wedge \\ u \leq z \wedge z \leq v \ \wedge \\ u \neq z \wedge z \neq v \end{pmatrix}$$

Datalog($\neq, \neg$) program for $\neg$PartialSucc$_n$:

$$\text{Succ}_n(u, v) \longleftarrow P_n(u), P_n(v), S_n(u, v), \text{STR}_{n-1}(u, v)$$

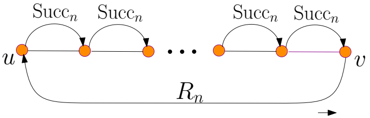
$$\text{NotPartialSucc}_n \longleftarrow \text{Succ}_n(u, v), X(u, v)$$

$$X(u, v) \longleftarrow P_n(z), u \leq z, z \leq v, u \neq z, z \neq v$$

SomeTotalR_n as a Datalog$(\neq, \neg)$ program

$$\text{SomeTotalR}_n := (\text{LO} \wedge \text{PartialSucc}_n) \rightarrow \exists u \exists v \; \text{RTotal}_n(u, v)$$

$\text{RTotal}_n(u, v) :=$



Datalog$(\neq, \neg)$ programs for $\text{RTotal}_n(u, v)$ and $\text{STR}_n(x, y)$:

$$\text{RTotal}_n(u, v) \longleftarrow R_n(v, u) \; \text{Total}_n(u, v)$$

$$\text{Total}_n(u, v) \longleftarrow \text{Succ}_n(u, v) \;\mid\; \text{Succ}_n(u, z), \text{Total}_n(z, v)$$

$$\text{STR}_n(x, y) \longleftarrow \text{NotLO}, \text{NotPartialSucc}_n,$$
$$x \leq u, v \leq y, \text{RTotal}_n(u, v)$$

Inexpressibility of SomeTotalR$_n$ in Π_{2n+1}
via showing $\mathcal{N}_{n,k} \Rrightarrow_{2n+1,k} \mathcal{M}_{n,k}$

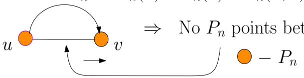
Proof approach

$$\text{SomeTotalR}_n := (\text{LO} \wedge \text{PartialSucc}_n) \to \exists u \exists v \ \text{RTotal}_n(u, v)$$

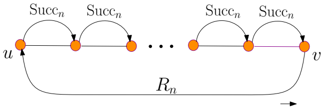
$\text{LO} :=$ " $\leq$ is a linear order"

$\text{PartialSucc}_n := \forall u \forall v$

$$\text{Succ}_n := P_n(u) \wedge P_n(v) \wedge S_n(u, v) \wedge \text{SomeTotalR}_{n-1}^{[u,v]}$$



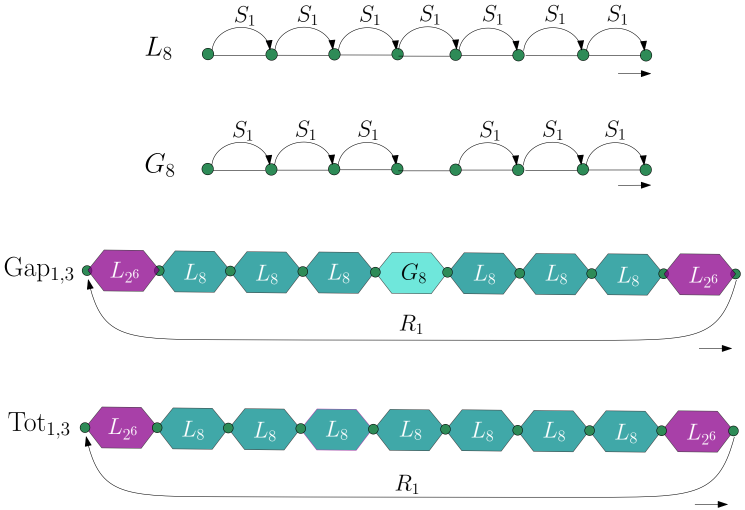
$\text{RTotal}_n(u, v) :=$



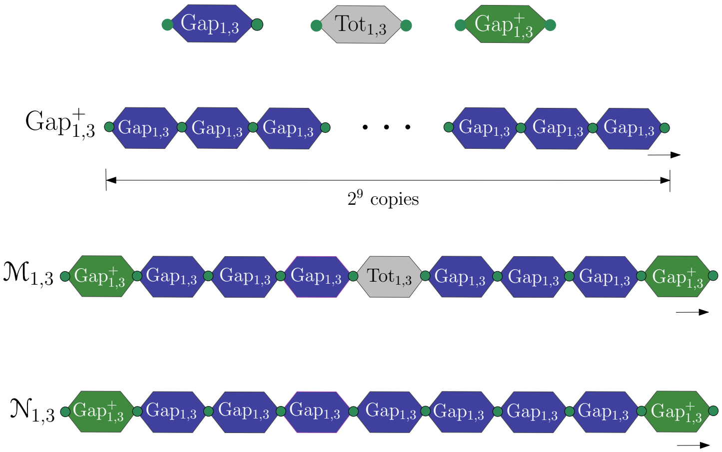
- For each k, we construct a model $\mathcal{M}_{n,k}$ and a non-model $\mathcal{N}_{n,k}$ of SomeTotalR$_n$ such that $\mathcal{N}_{n,k} \Rrightarrow_{2n+1,k} \mathcal{M}_{n,k}$ holds.
- Then for every $\Pi_{2n+1,k}$ sentence θ, we have $\mathcal{M}_{n,k} \models \theta \to \mathcal{N}_{n,k} \models \theta$; then $\theta \not\leftrightarrow$ SomeTotalR$_n$.

Construction of $\mathcal{M}_{n,k}$ and $\mathcal{N}_{n,k}$
(Illustrated for $k = 3$)

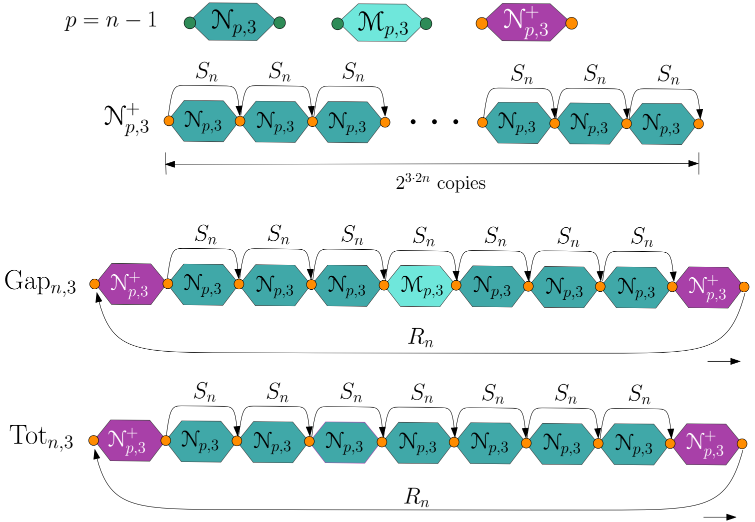
$\mathcal{M}_{1,3}$ and $\mathcal{N}_{1,3}$



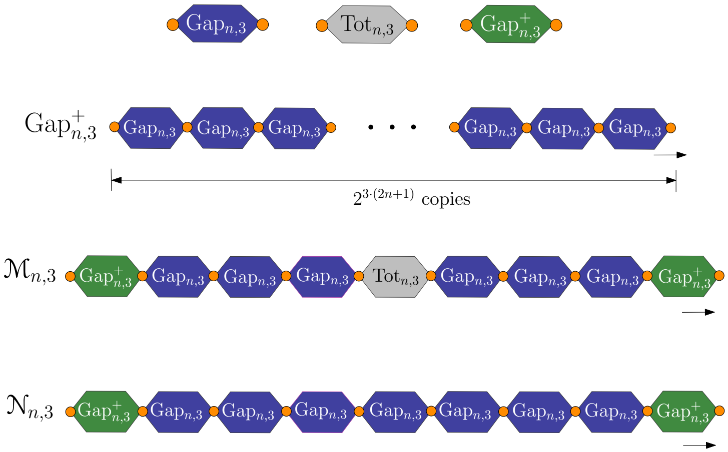
$\mathcal{M}_{1,3}$ and $\mathcal{N}_{1,3}$



$\mathcal{M}_{n,3}$ and $\mathcal{N}_{n,3}$



$\mathcal{M}_{n,3}$ and $\mathcal{N}_{n,3}$



Inexpressibility of $\mathsf{SomeTotalR}_n$ in Π_{2n+1}

All of the following can be shown analogously to the corresponding statements for Tait's sentence.

- The sentence $\mathsf{SomeTotalR}_n$ is not extension closed in the infinite.
- $\mathcal{M}_{n,k} \models \mathsf{SomeTotalR}_n$ but $\mathcal{N}_{n,k} \not\models \mathsf{SomeTotalR}_n$.
- $\mathsf{Tot}_{n,k} \Rrightarrow_{2n,k} \mathsf{Gap}_{n,k}$ and $\mathcal{N}_{n,k} \Rrightarrow_{2n+1,k} \mathcal{M}_{n,k}$.
- Then every $\Pi_{2n+1,k}$ sentence true in $\mathcal{M}_{n,k}$ is also true in $\mathcal{N}_{n,k}$; whereby $\mathsf{SomeTotalR}_n$ cannot be equivalent to a $\Pi_{2n+1,k}$ sentence.

Conclusion

Main results revisited

Theorem

Tait's counterexample is a Σ_3 FO sentence that is extension preserved over all finite structures, but is not equivalent over this class to any Π_3 sentence. Further, the counterexample can be expressed in Datalog($\neq, \neg$).

Theorem

For every n, there is a vocabulary σ_n and an FO(σ_n) Σ_{2n+1} sentence φ_n that is extension closed over all finite structures, but that is not equivalent over this class to any Π_{2n+1} sentence. Further, φ_n can be expressed in Datalog($\neq, \neg$).

Main results revisited

Theorem

Tait's counterexample is a Σ_3 FO sentence that is extension preserved over all finite structures, but is not equivalent over this class to any Π_3 sentence. Further, the counterexample can be expressed in Datalog$(\neq, \neg)$.

Theorem

No prefix class of FO is expressive enough to capture:

- Extension closed FO properties in the finite
- FO $\cap$ Datalog$(\neq, \neg)$ queries in the finite

Future directions

- The sentence SomeTotalR_n is over a vocabulary σ_n that grows with n.
- Further, σ_n can be seen as the vocabulary of ordered vertex colored and edge colored graphs.

Question 1.

Is there a fixed (finite) vocabulary σ^* such that prefix classes fail to capture extension preserved FO properties of finite σ^*-structures?

Question 2.

Do prefix classes fail to capture extension preserved FO properties of undirected graphs (possibly vertex colored)?

Future directions

- The sentence SomeTotalR$_n$ is over a vocabulary σ_n that grows with n.
- Further, σ_n can be seen as the vocabulary of ordered vertex colored and edge colored graphs.

Question 1. (Resolved: Yes! $|\sigma^*| \leq 4$)

Is there a fixed (finite) vocabulary σ^* such that prefix classes fail to capture extension preserved FO properties of finite σ^*-structures?

Question 2. (Not resolved yet)

Do prefix classes fail to capture extension preserved FO properties of undirected graphs (possibly vertex colored)?

Thank you!