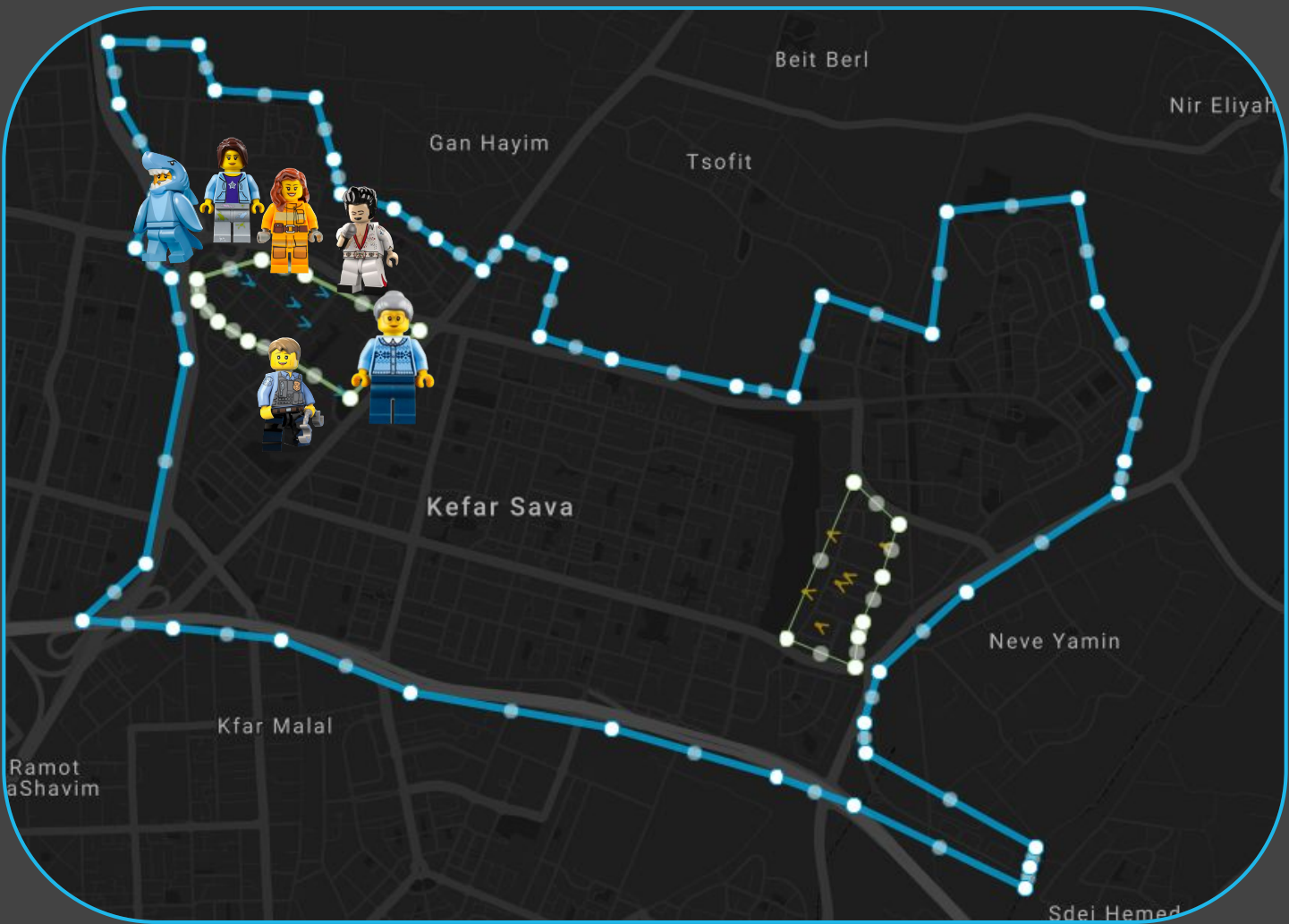
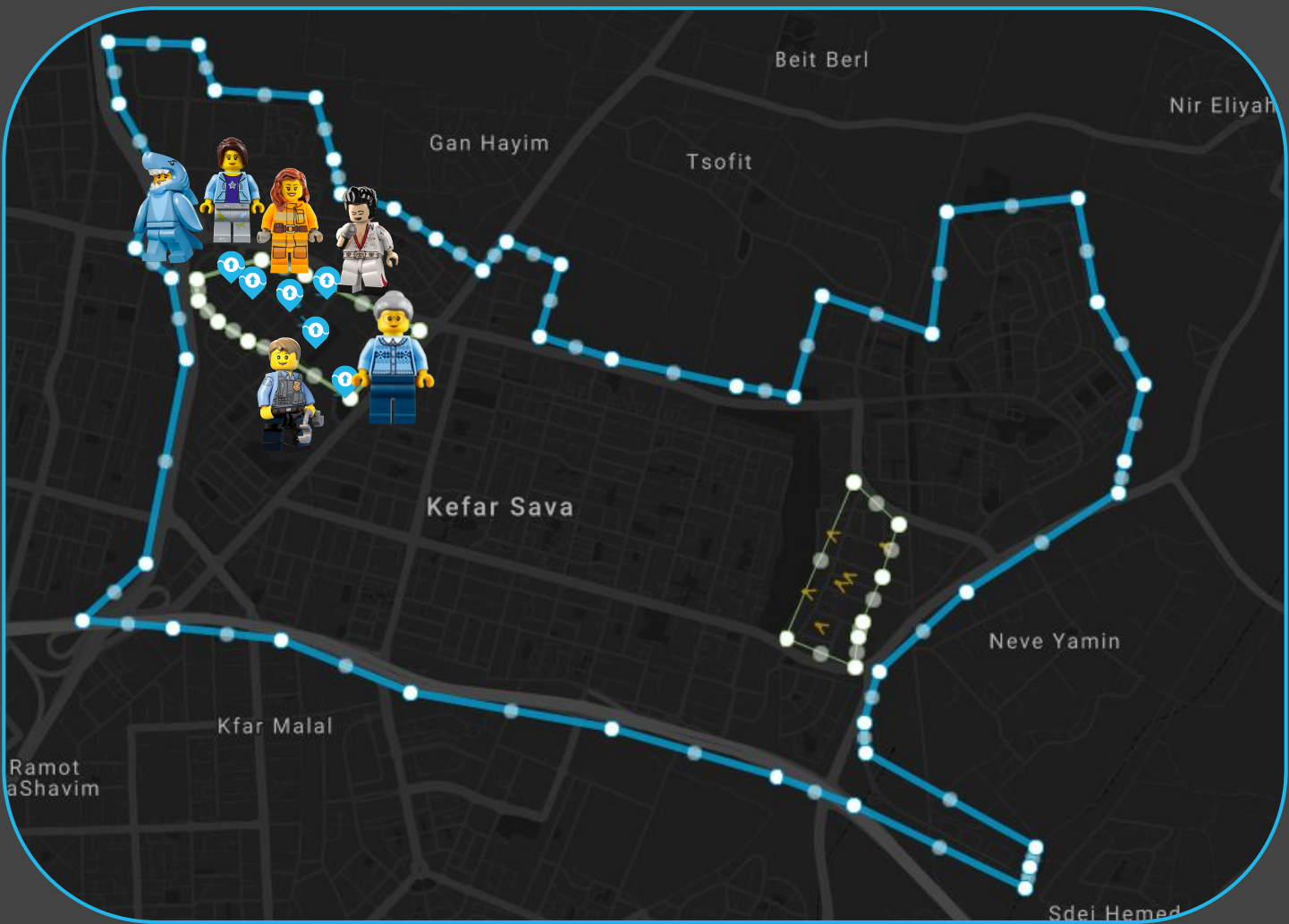


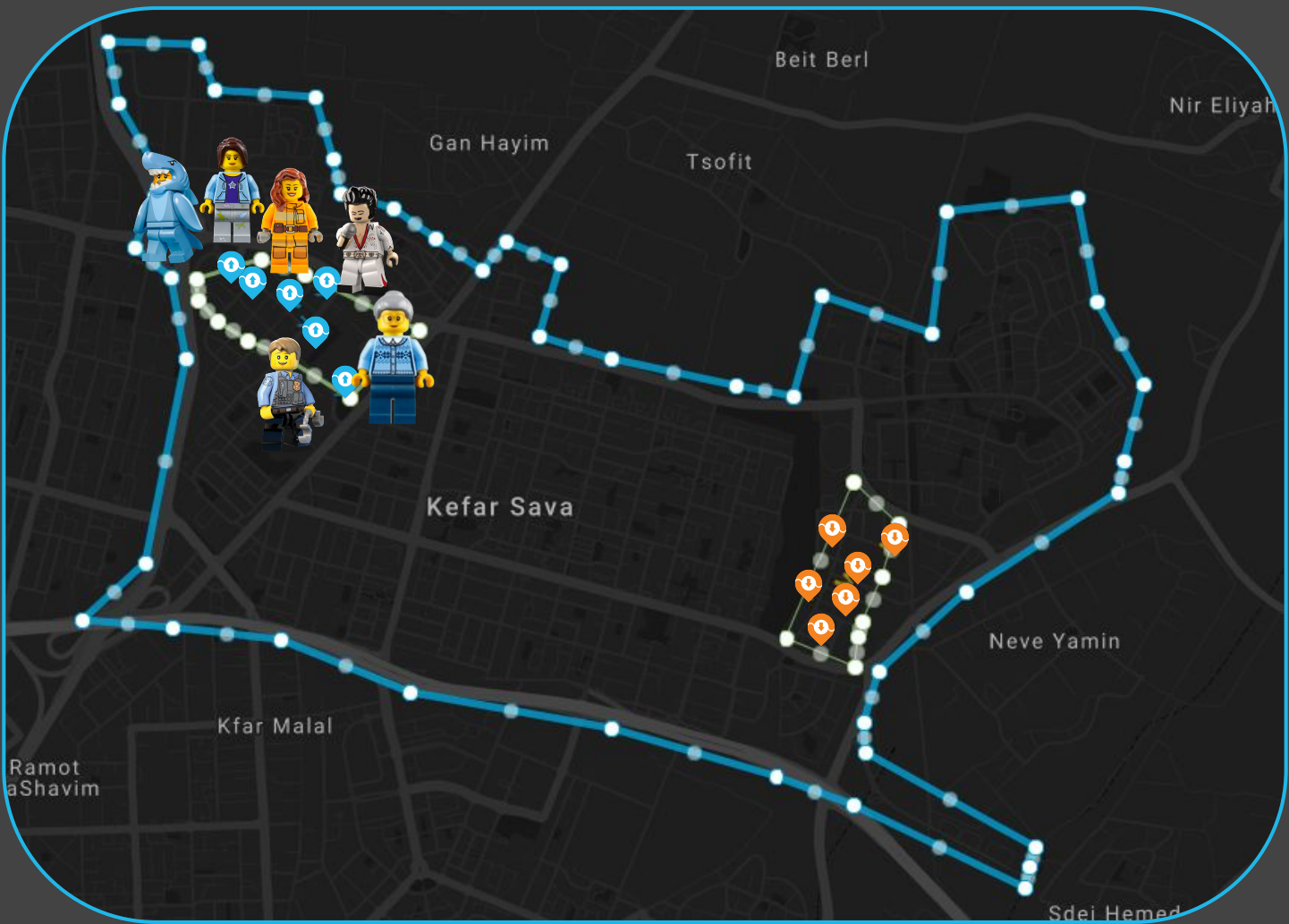
Sequence Alignment for Ride Sharing

Dalya Gartzman













**~7.5 million
arrangements!!**

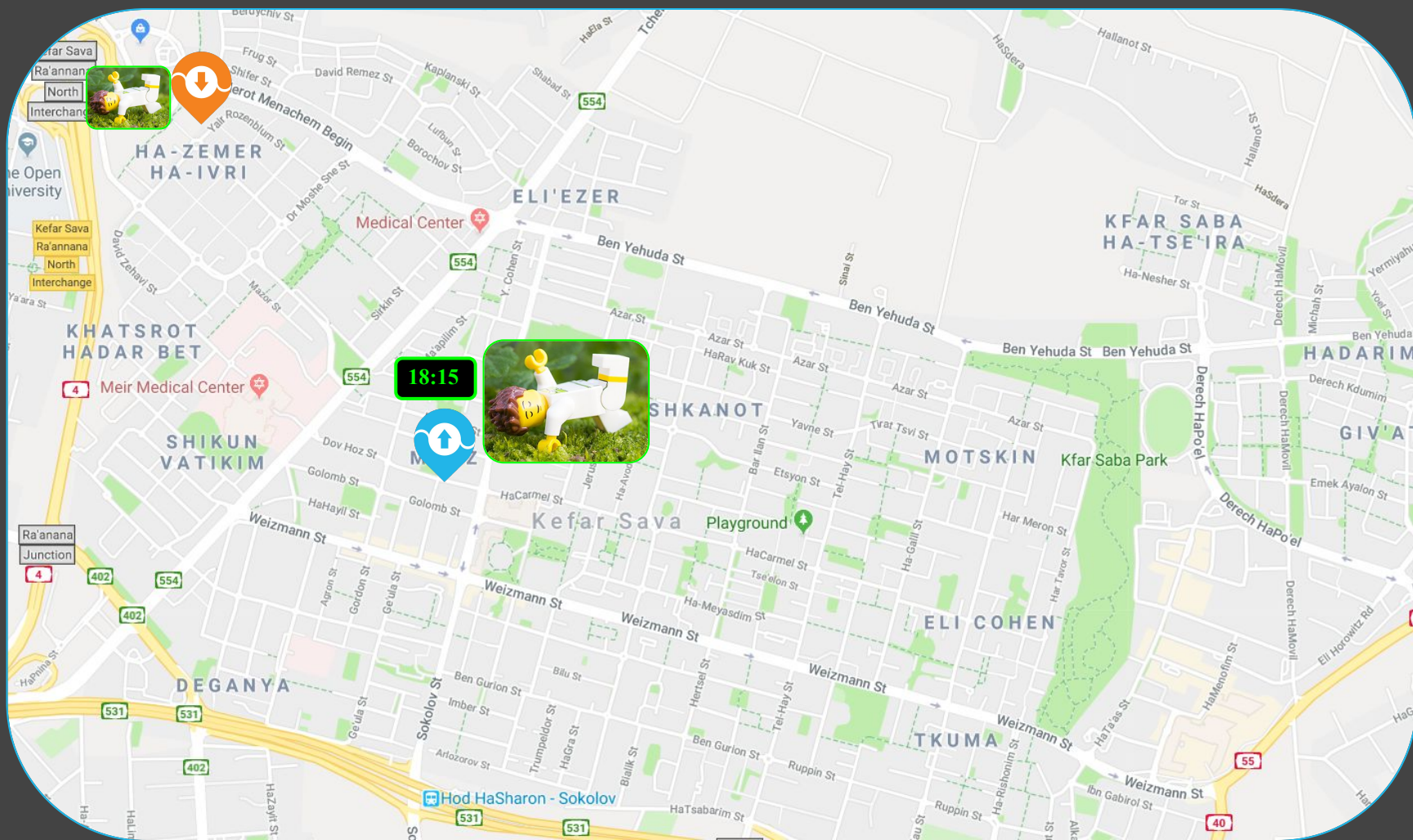


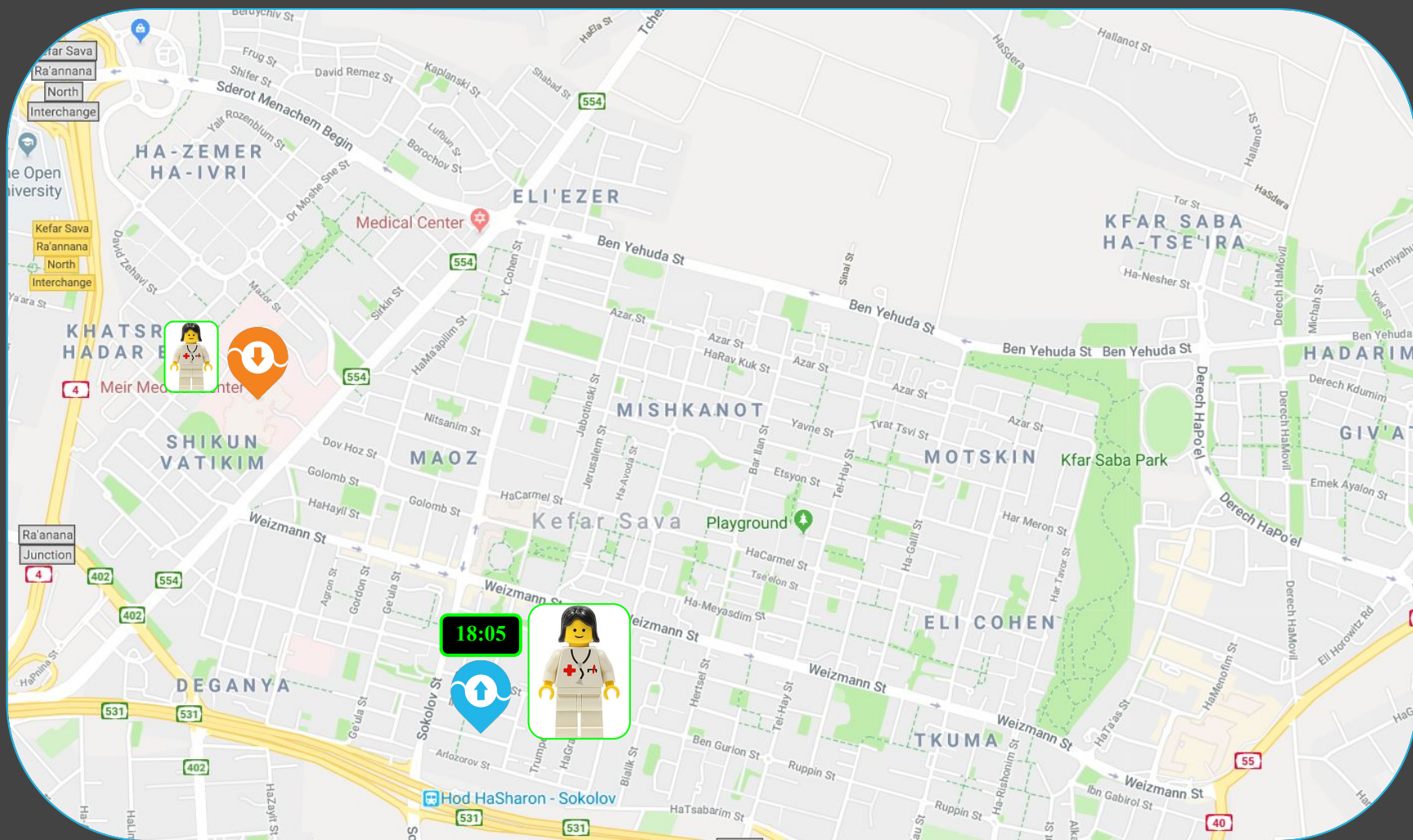
Given n riders who want to
get from A_i to B_i , find the
best Taxi Ride for everyone.

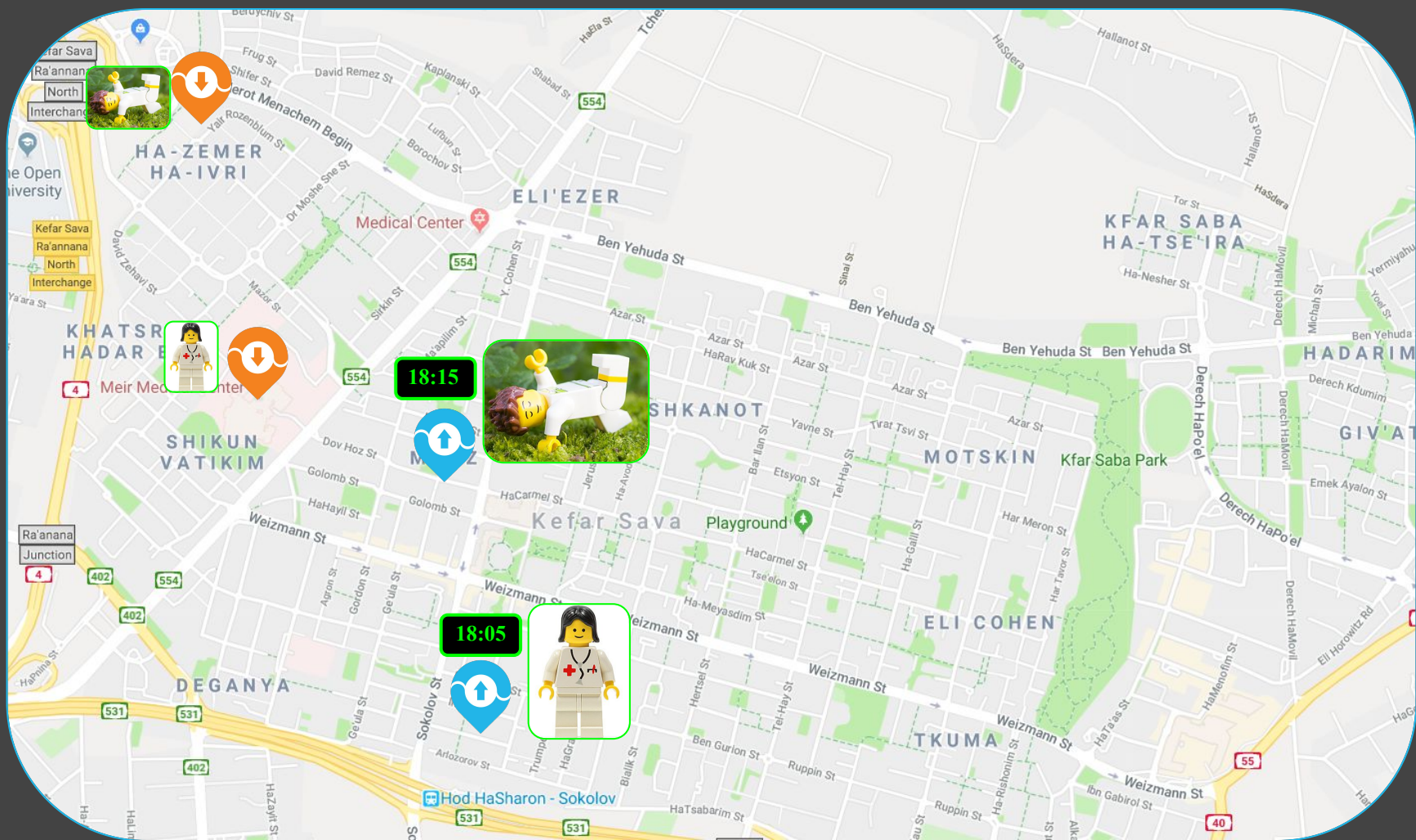
Given n riders who want to
get from A_i to B_i , find the
best Taxi Ride for everyone.

Given n riders who want to
get from A_i to B_i , find the
best Taxi Ride for everyone.

Given n riders who want to
get from A_i to B_i , find the
best Taxi Ride for everyone.







Ps

1



Hod HaSharon - Sokolov

PsPr

2



1



Hod HaSharon - Sokolov

P_sP_rD_s

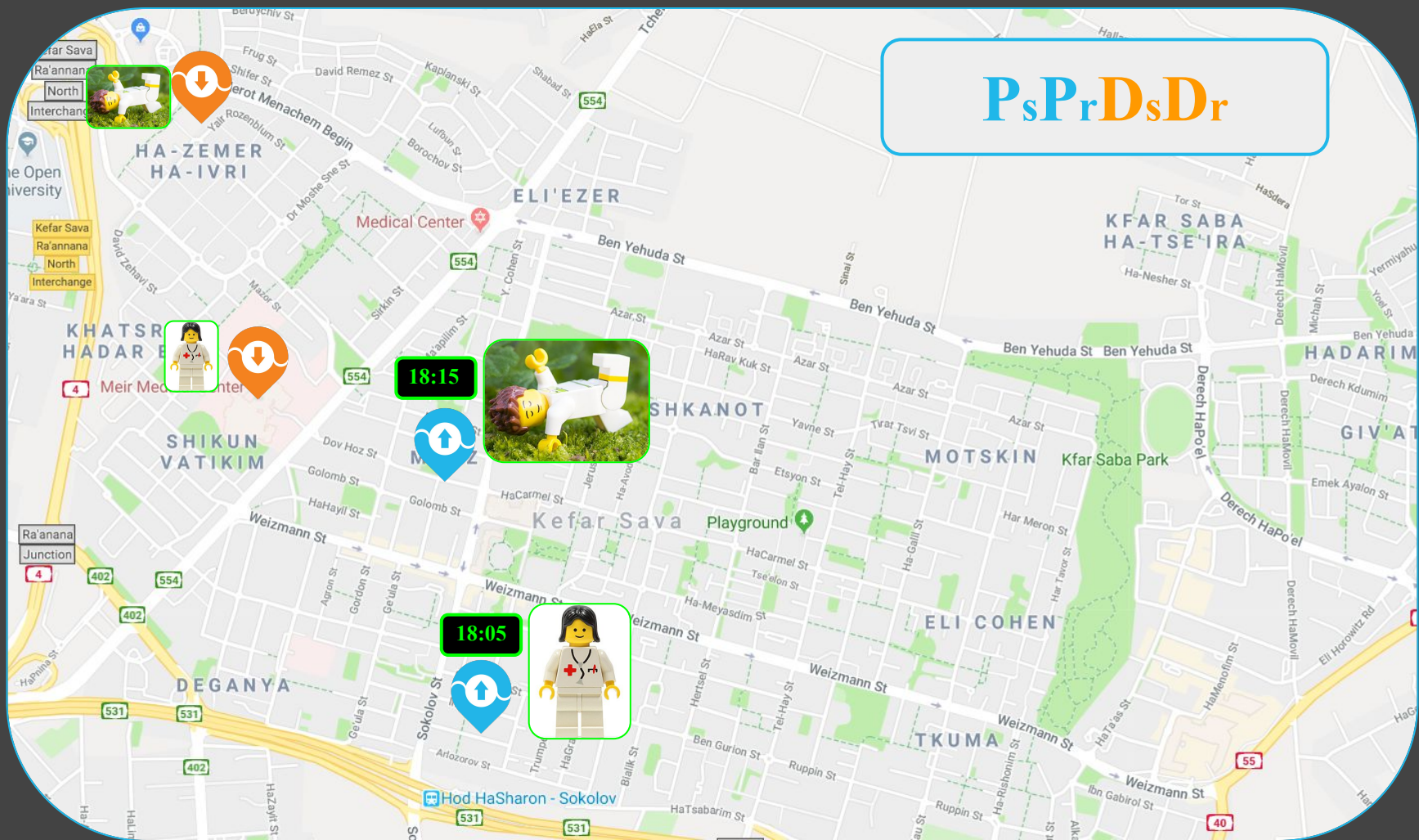


PsPrDsDr



“BEST” = ?

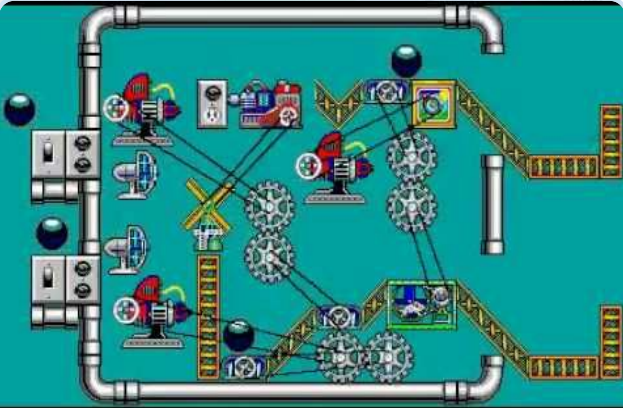
PsPrDsDr



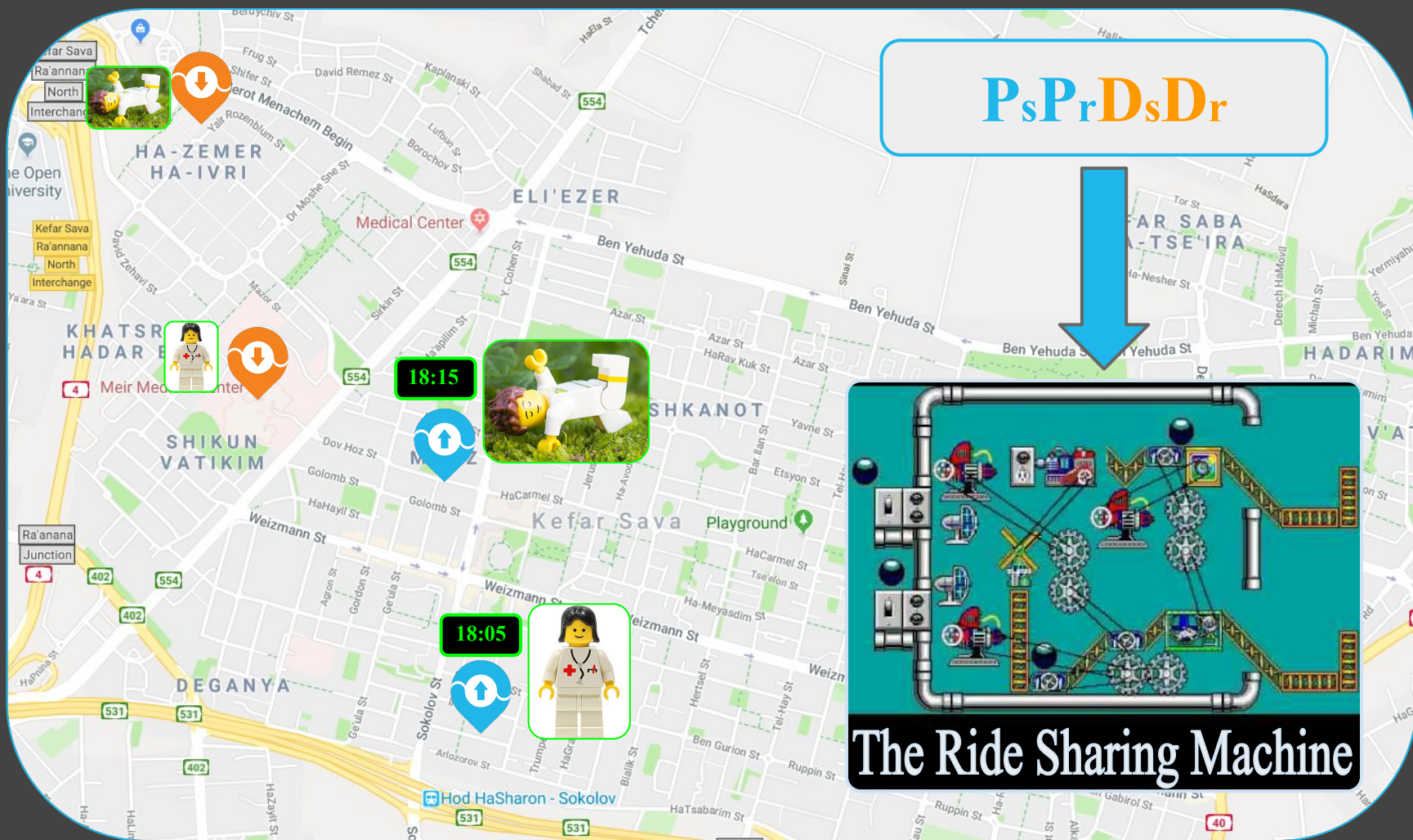
18:15

18:05

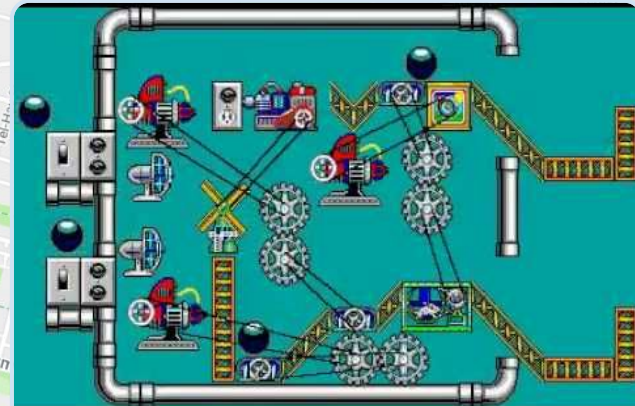
P_sP_rD_sD_r



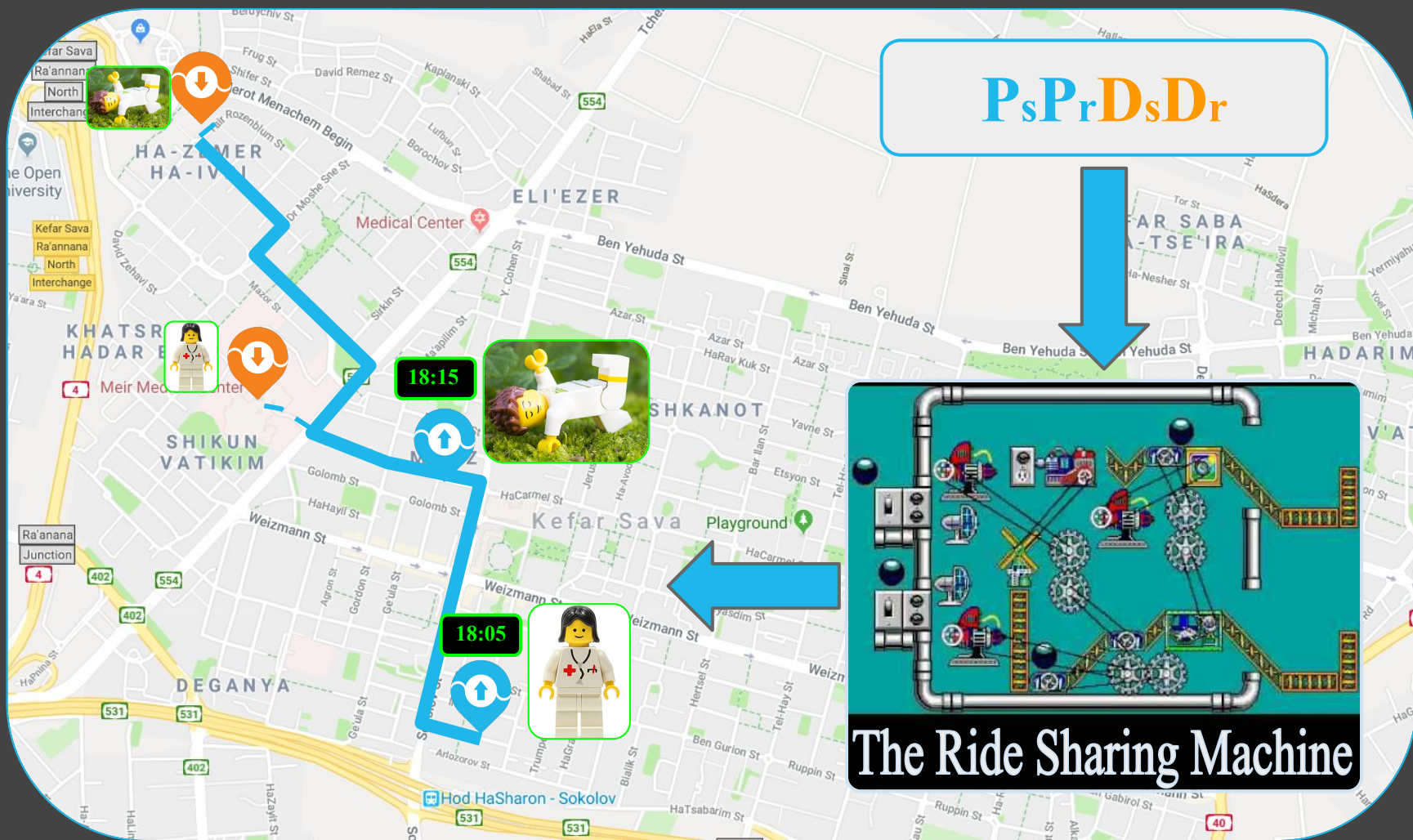
The Ride Sharing Machine



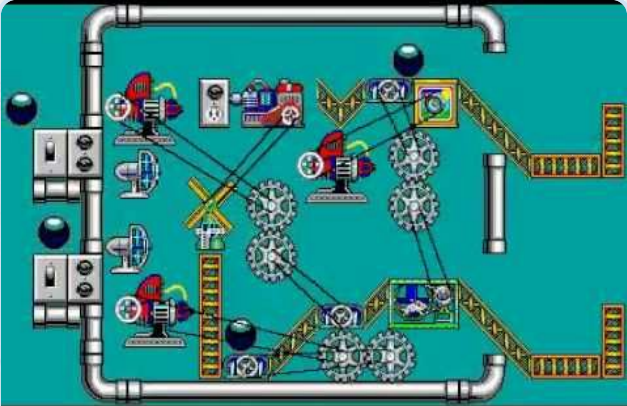
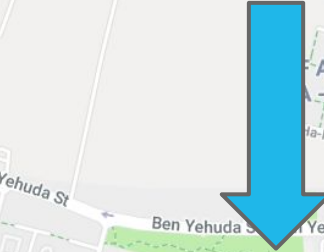
PsPrDsDr



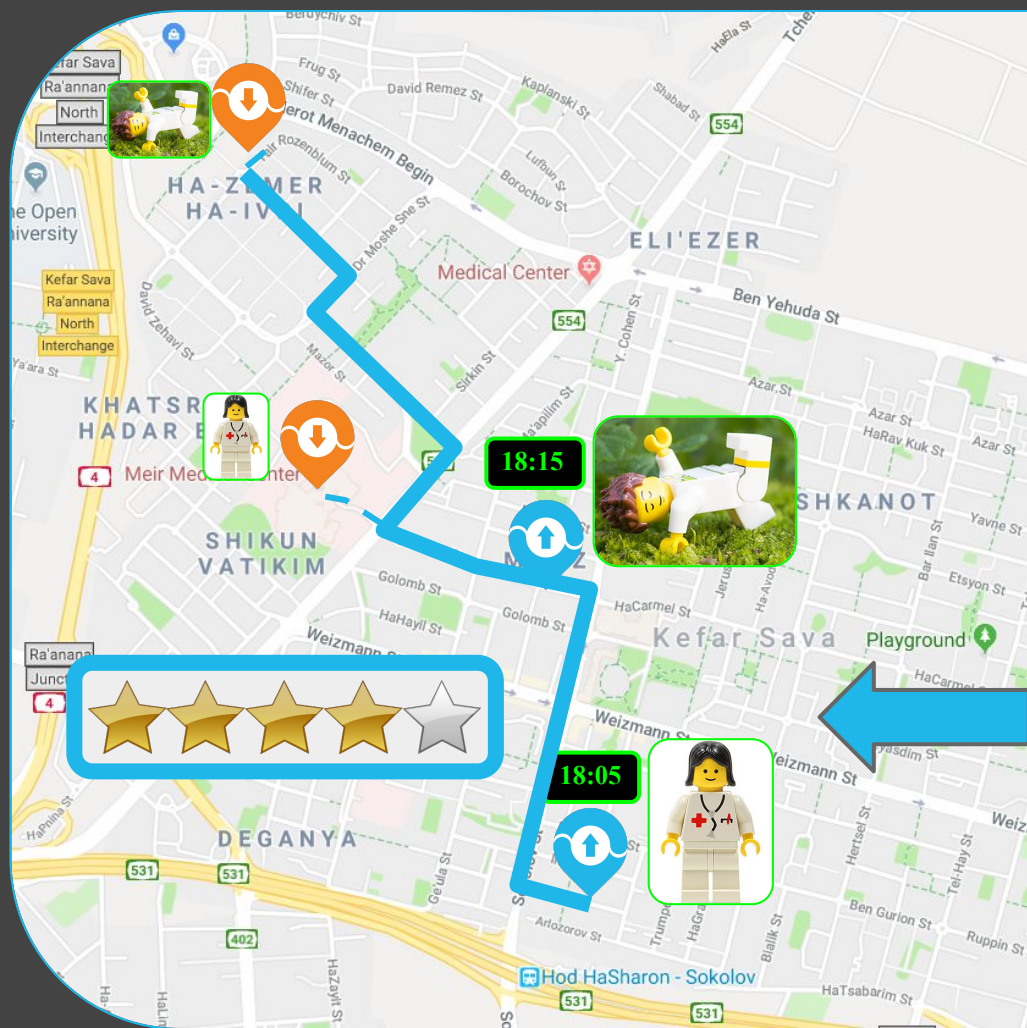
The Ride Sharing Machine

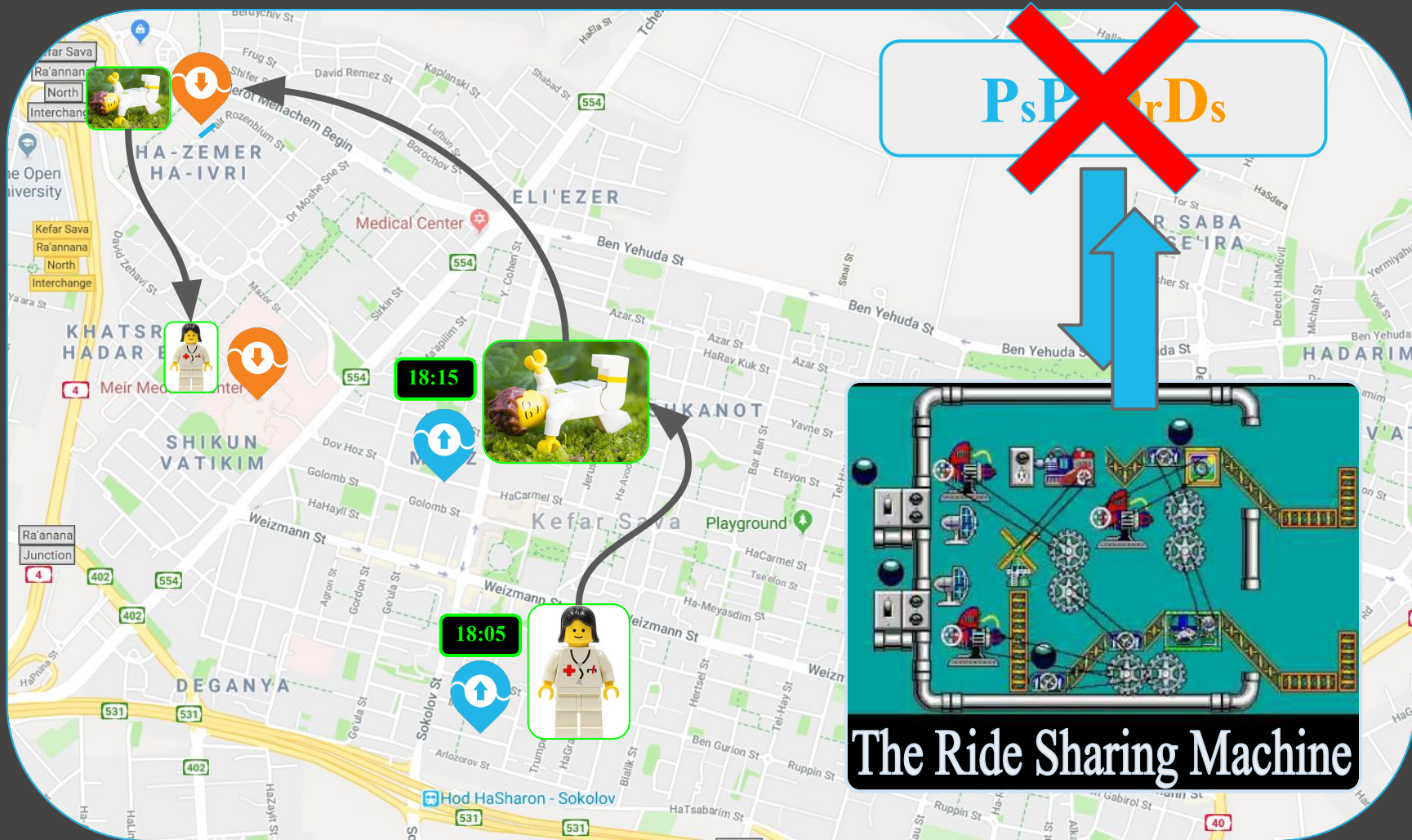


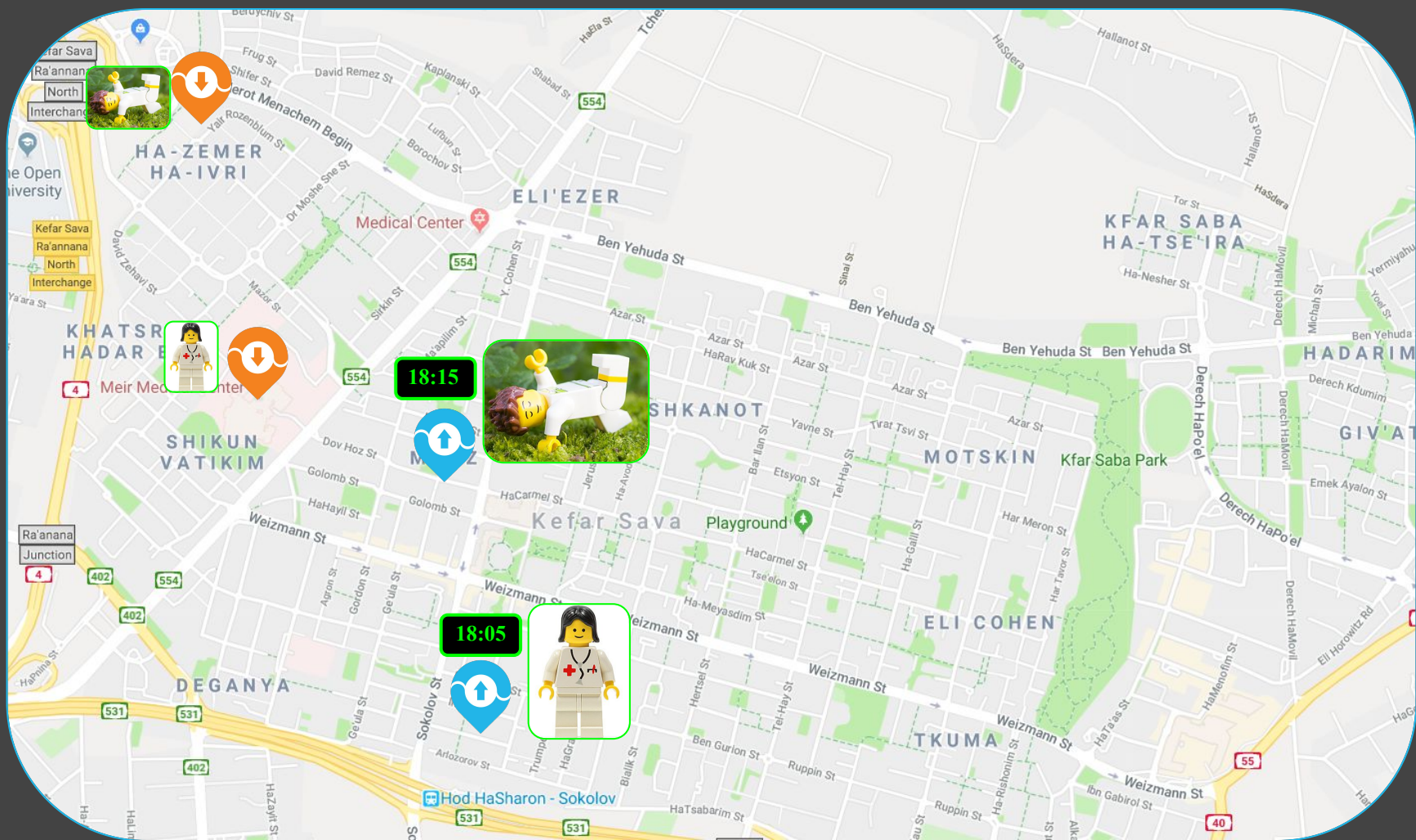
P_sP_rD_sD_r



The Ride Sharing Machine



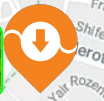


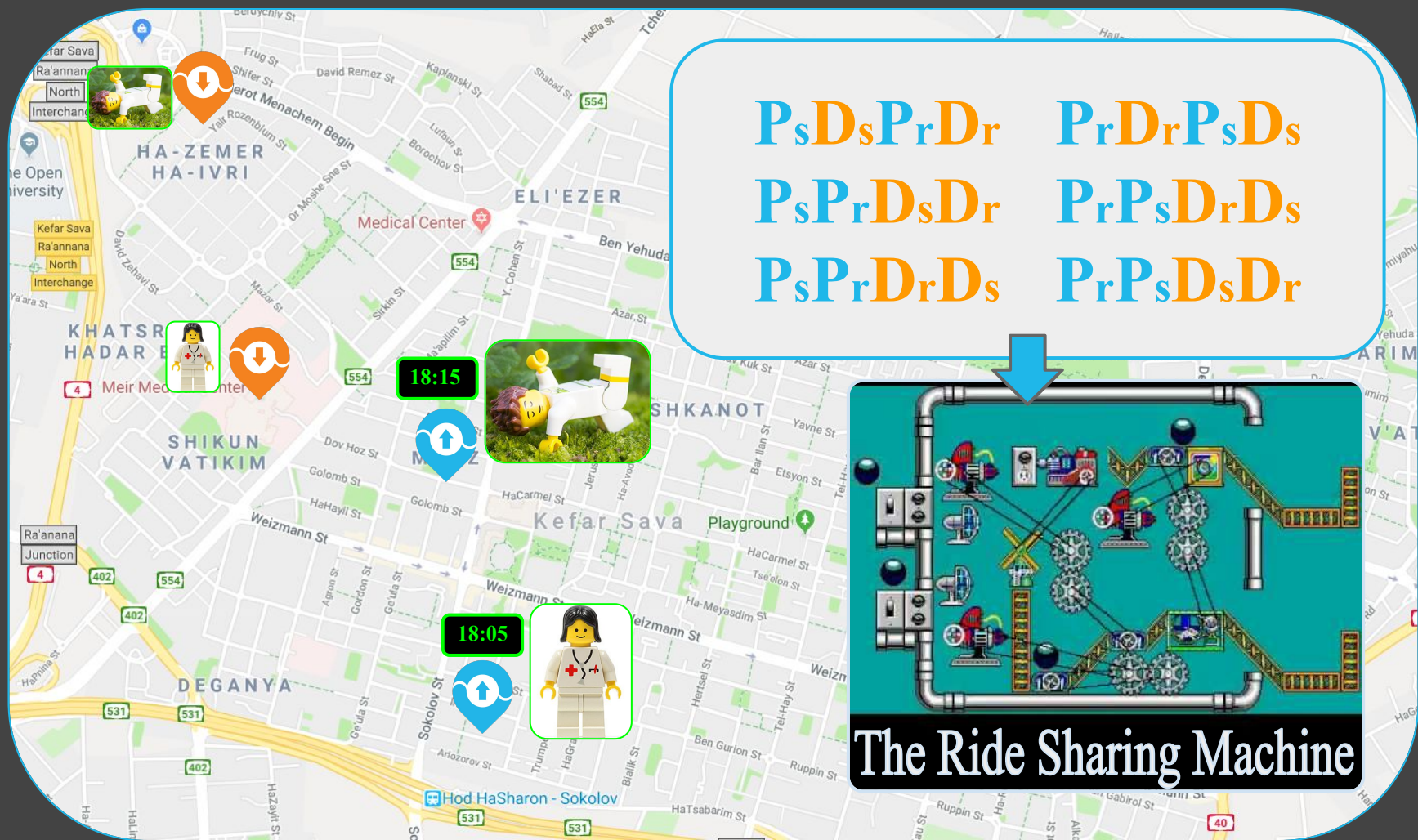


PrDrPsDs

PrPsDrDs

PrPsDsDr





PsDsPrDr

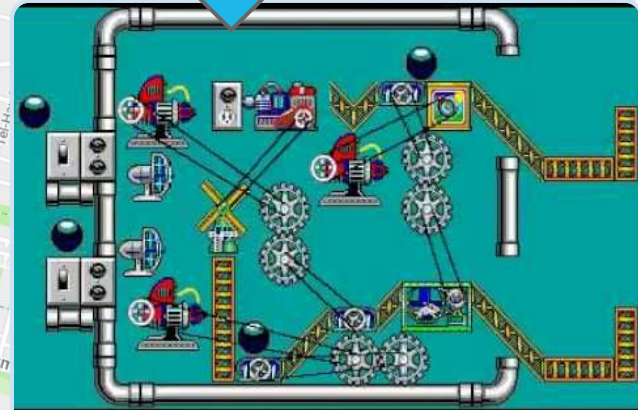
PrDrPsDs

PsPrDsDr

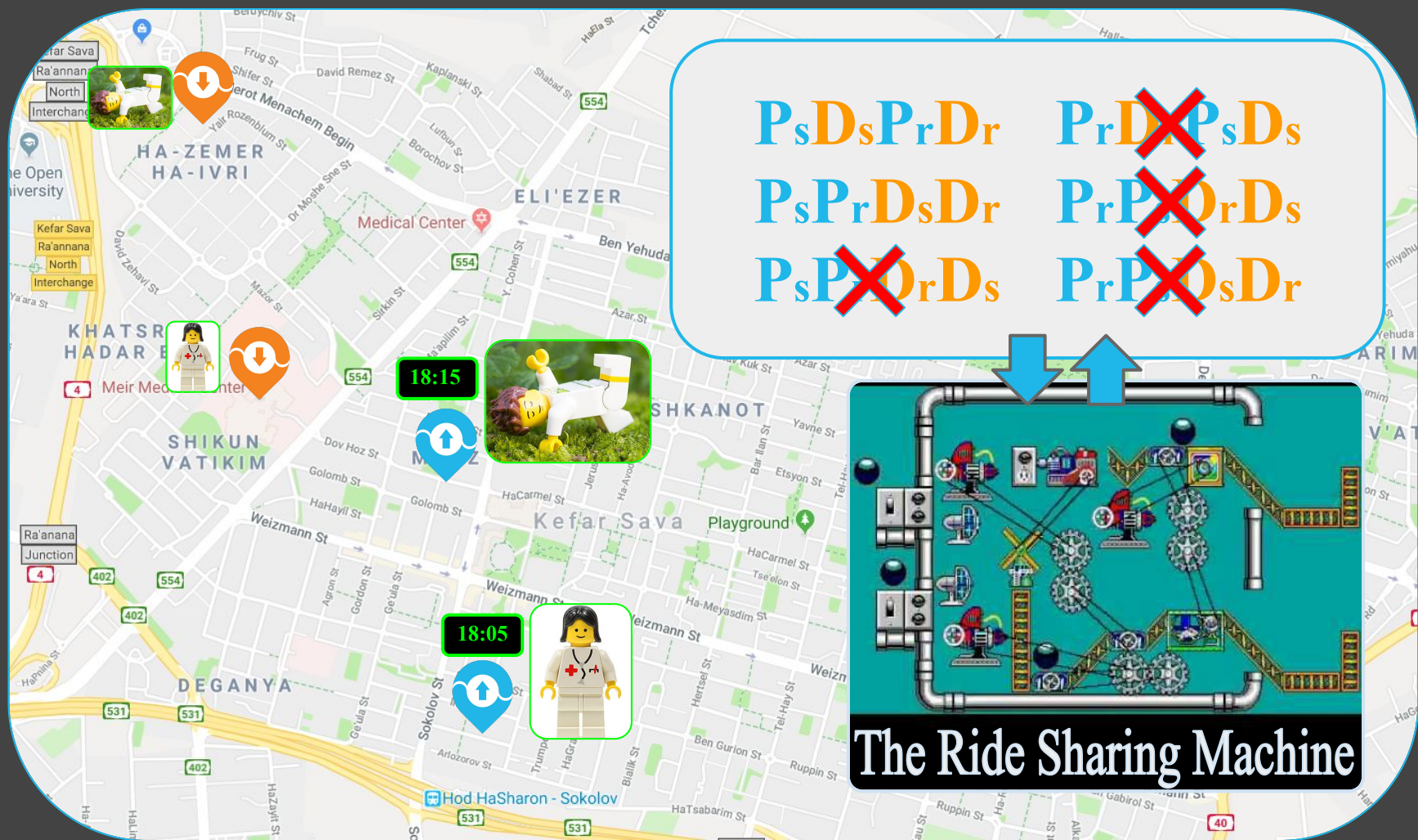
PrPsDrDs

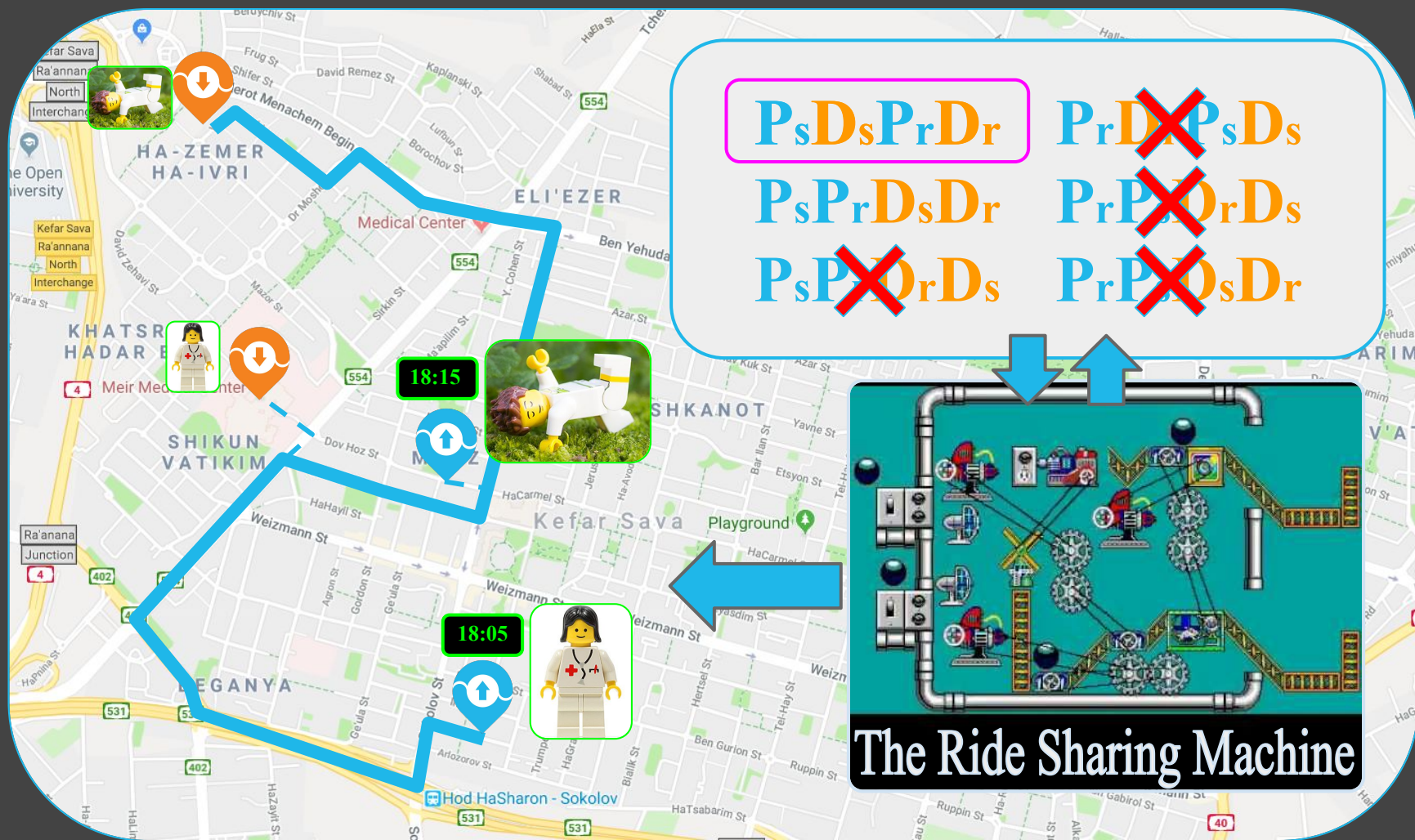
PsPrDrDs

PrPsDsDr



The Ride Sharing Machine





$P_s D_s P_r D_r$

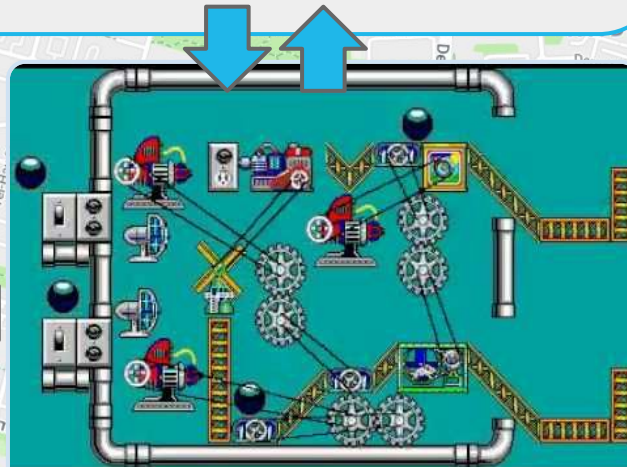
$P_r D_r P_s D_s$

$P_s P_r D_s D_r$

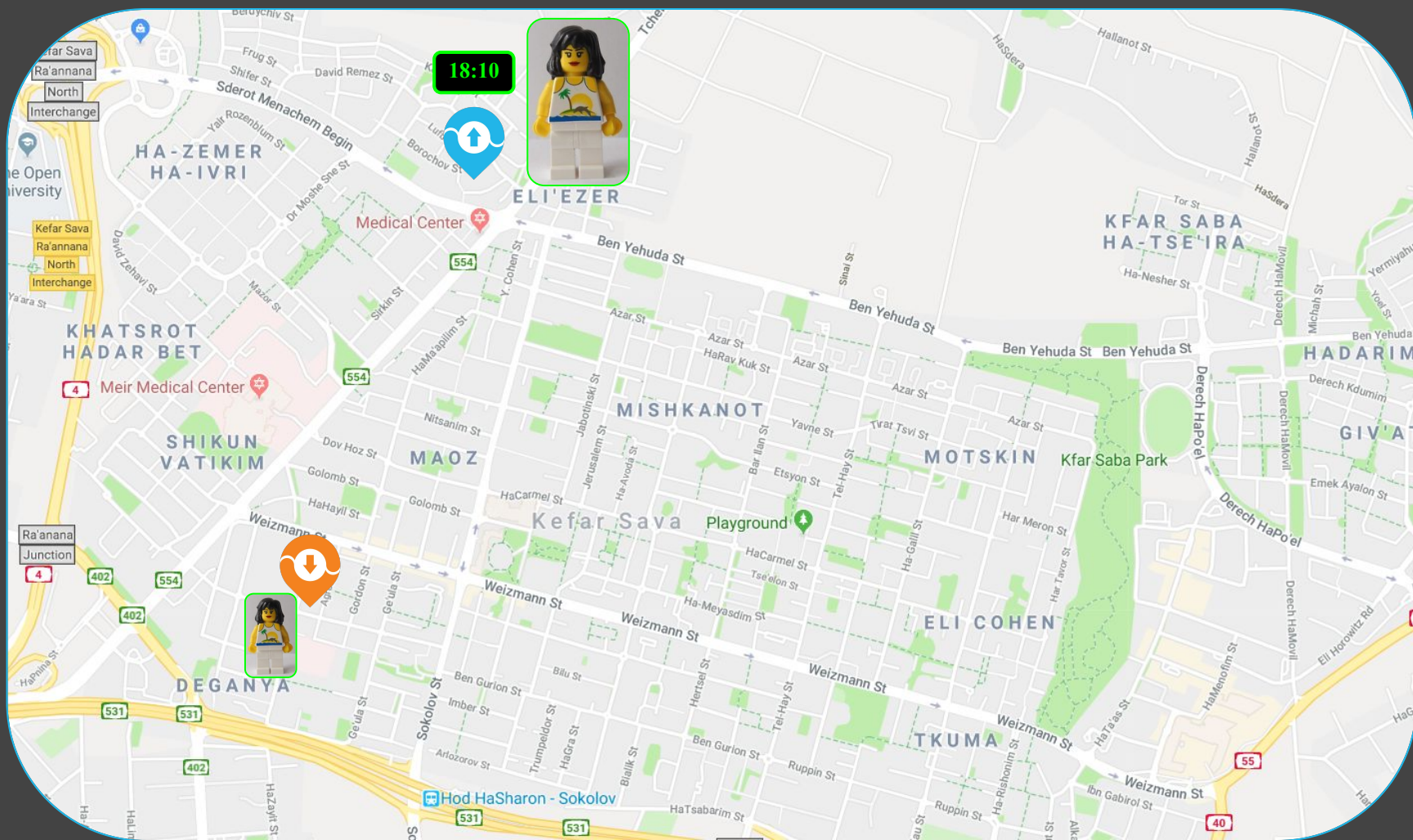
$P_r P_r D_r D_s$

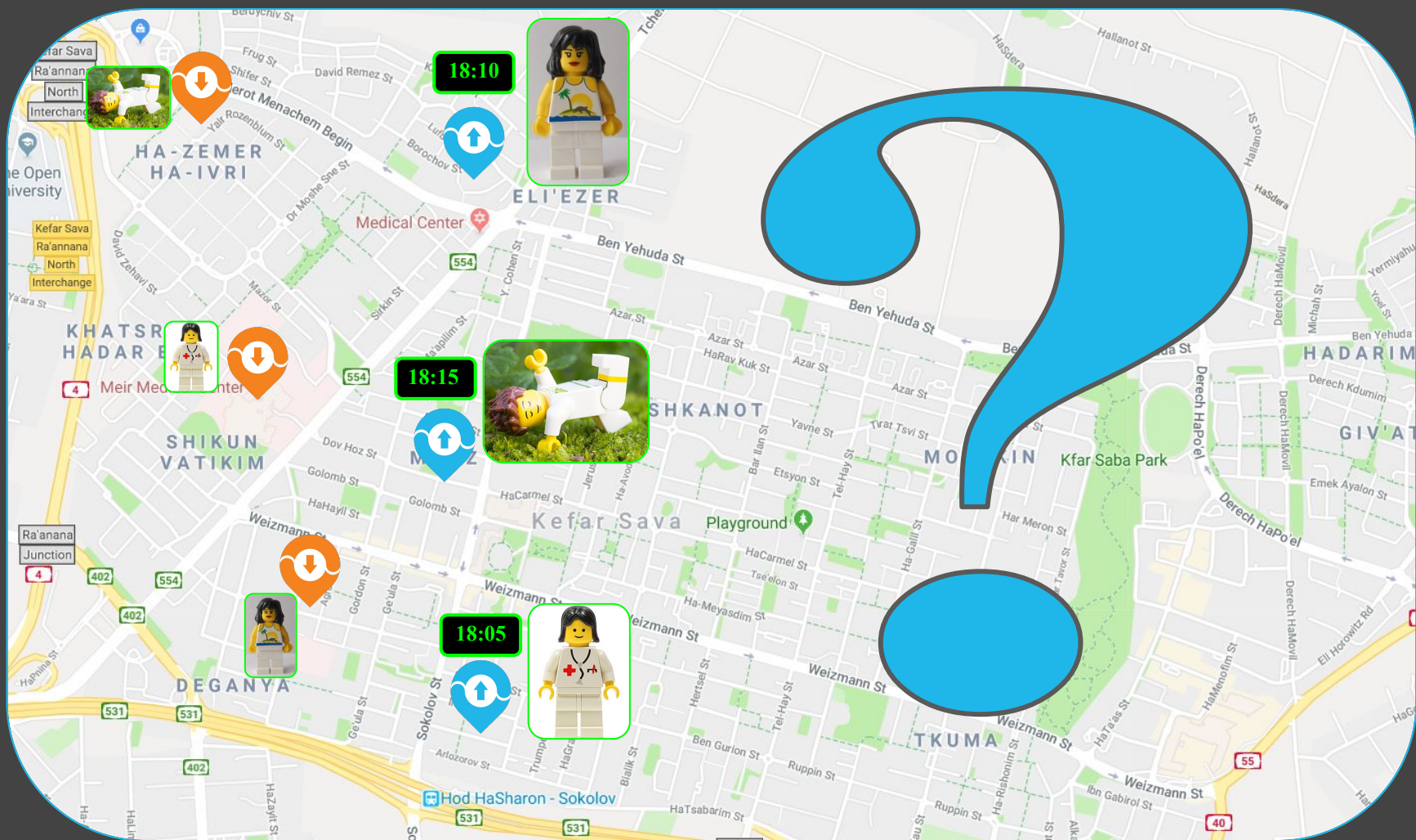
$P_s P_r D_r D_s$

$P_r P_r D_s D_r$



The Ride Sharing Machine





How many permutations?

How many permutations?

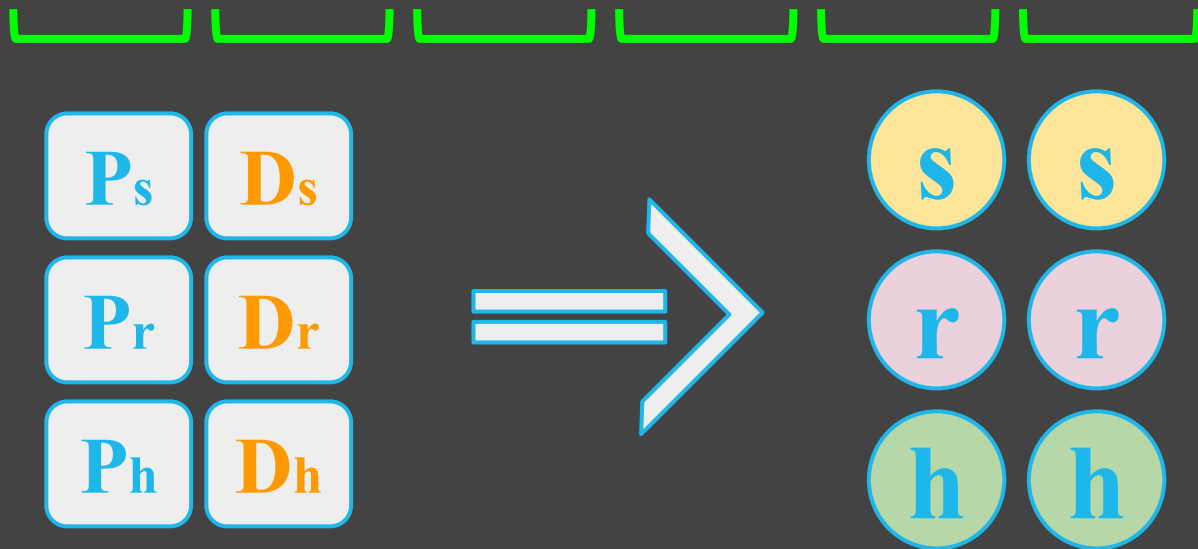


How many permutations?

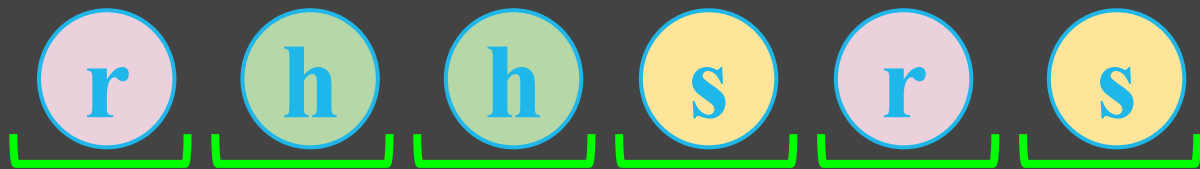


P_s	D_s
P_r	D_r
P_h	D_h

How many permutations?

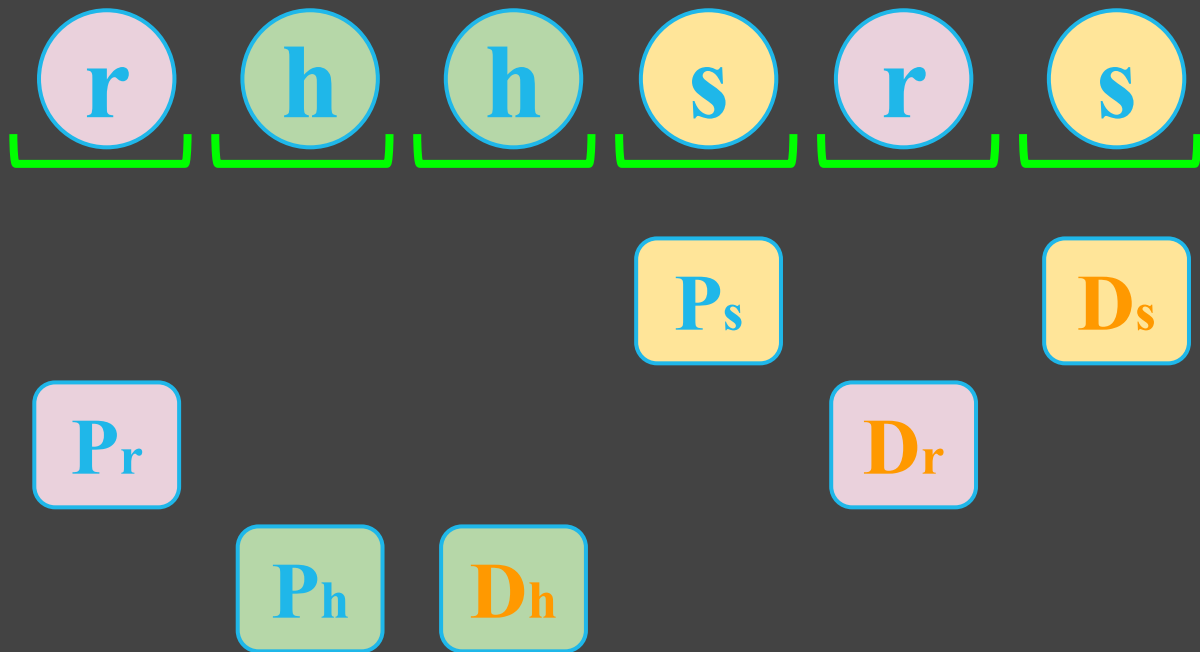


How many permutations?



P_s	D_s
P_r	D_r
P_h	D_h

How many permutations?



How many permutations?

The diagram illustrates the calculation of permutations for the sequence **r h h s r s**. The letters are grouped into three pairs, each with a corresponding permutation box below it:

- r** (pink circle) and **P_r** (pink box)
- h** (green circle) and **P_h** (green box)
- h** (green circle) and **D_h** (orange box)
- s** (yellow circle) and **P_s** (yellow box)
- r** (pink circle) and **D_r** (pink box)
- s** (yellow circle) and **D_s** (orange box)

The final result is shown in a white rounded rectangle:

$$\frac{6!}{(2!)^3}$$

How many permutations?

$$\frac{(2n)!}{2^n}$$

How many permutations?


$$\frac{(2n)!}{2^n}$$

n	1	2	3	4	5	6
$2^{-n} (2n)!$	1	6	90	2520	113400	7484400

How many permutations?



$$\frac{(2n)!}{2^n}$$



	2	3	4	5	6	
$2^{-n} (2n)!$	1	6	90	2520	113400	7484400

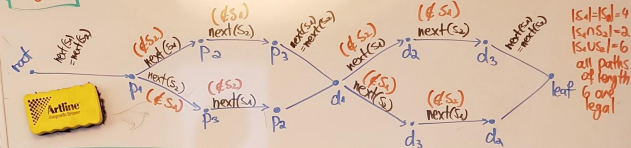


Reduce number of
permutations!

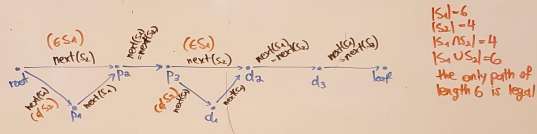
DAG+DP? MODLE

$S_1 = p_1 p_2 d_1 d_2 d_3$ $S_2 = p_1 p_2 d_1 d_3$

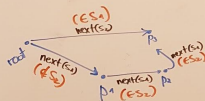
vertices operations insights



$S_1 = p_1 p_2 p_3 d_1 d_2 d_3$ $S_2 = p_1 p_2 d_1 d_3$ $[S_2 \subset S_1]$



$S_1 = p_1 p_2 p_3 d_1 d_2 d_3$ $S_2 = p_1 p_2 d_1 d_3$ $[S_2 \not\subset S_1]$



$|S_1| = |S_2| \Rightarrow \text{JoinAndAlignSequences}$
 $[|S_1| = |S_2| = n-2, |S_1| = |S_2| = n-2]$

$|S_1| = |S_2| + 2 \Rightarrow \text{IsSubstring}$

given $\{d_1, \dots, d_k\}$, $|d_1| = k_1, |d_2| = k_2, \dots, |d_k| = k_k, |d_k| = k_k$

- all sequences = $[s \text{ for } s \text{ in JoinAndAlignSequences}(S_1, S_2) \text{ for } s \in I_1, p \in I_2]$
- for $i = 3 \dots k$: all sequences = $[s \text{ for } s \text{ in all sequences if } \exists d \in I, \text{ st. } \text{IsSubstring}(s, d)]$

DNA Sequence Alignment



DNA Sequence Alignment



Sequence1	TCAGGATGAACG
Sequence2	ATCACGATGAACC
Sequence3	TCAGGATTGAACCGC
Sequence4	ATCACGAATGAATCC

DNA Sequence Alignment



Sequence1	-TCAGGA-TGAAC-G-
Sequence2	ATCACGA-TGAACC--
Sequence3	-TCAGGATTGAACCGC
Sequence4	ATCACGAATGAATCC-

DNA Sequence Alignment



Sequence1	-TCAGGA-TGAAC-G-
Sequence2	ATCACGA-TGAACC--
Sequence3	-TCAGGATTGAACCGC
Sequence4	ATCACGAATGAATCC-

DNA Sequence Alignment



Sequence1	-TCAGGA-TGAAC-G-
Sequence2	ATCACGA-TGAACC--
Sequence3	-TCAGGATTGAACCGC
Sequence4	ATCACGAATGAATCC-

Our Sequence Alignment

Sequence1 P_h P_s D_h - - D_s

Sequence2 P_h - D_h P_r D_r -

Our Sequence Alignment

Sequence1 **P_h** **P_s** **D_h** - - **D_s**

Sequence2 **P_h** - **D_h** **P_r** **D_r** -

Result **P_h** **P_s** **D_h** **P_r** **D_r** **D_s**

Our Sequence Alignment

Sequence1 P_h P_s D_h - - D_s

Sequence2 P_h - D_h P_r D_r -

Result P_h P_s D_h P_r D_r D_s



Our Sequence Alignment

Sequence1	P_h	P_s	D_h	-	-	D_s
Sequence2	P_h	-	D_h	P_r	D_r	-
Result	P_h	P_s	D_h	P_r	D_r	D_s



Our Sequence Alignment

Sequence1

P_h P_s D_h - - D_s

Sequence2

P_h - D_h P_r D_r -

Result

P_h P_s D_h P_r D_r D_s



We Can Do It!



Sequence Alignment for Ride Sharing



$P_h P_s D_h D_s$



$P_h D_h P_r D_r$

Sequence Alignment for Ride Sharing



P_h P_s D_h D_s



P_h D_h P_r D_r

Sequence Alignment for Ride Sharing



P_h

P_s

D_h

D_s



P_h

D_h

P_r

D_r

Sequence Alignment for Ride Sharing



$P_h P_s D_h D_s$



$P_h D_h P_r D_r$

Sequence1

$P_h P_s D_h - - D_s$



+

Sequence2

$P_h - D_h P_r D_r -$



=

Result

$P_h P_s D_h P_r D_r D_s$



$P_h P_s D_h D_s$

$P_h D_h P_r D_r$

$P_h P_s D_h D_s$

$P_h D_h P_r D_r$

**mutual
subsequence**

$P_h P_s D_h D_s$

$P_h D_h P_r D_r$

mutual
subsequence

START

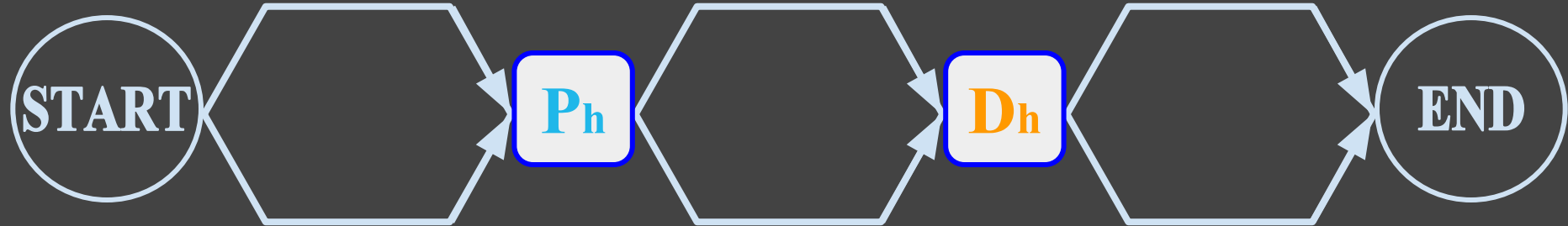
P_h

D_h

END

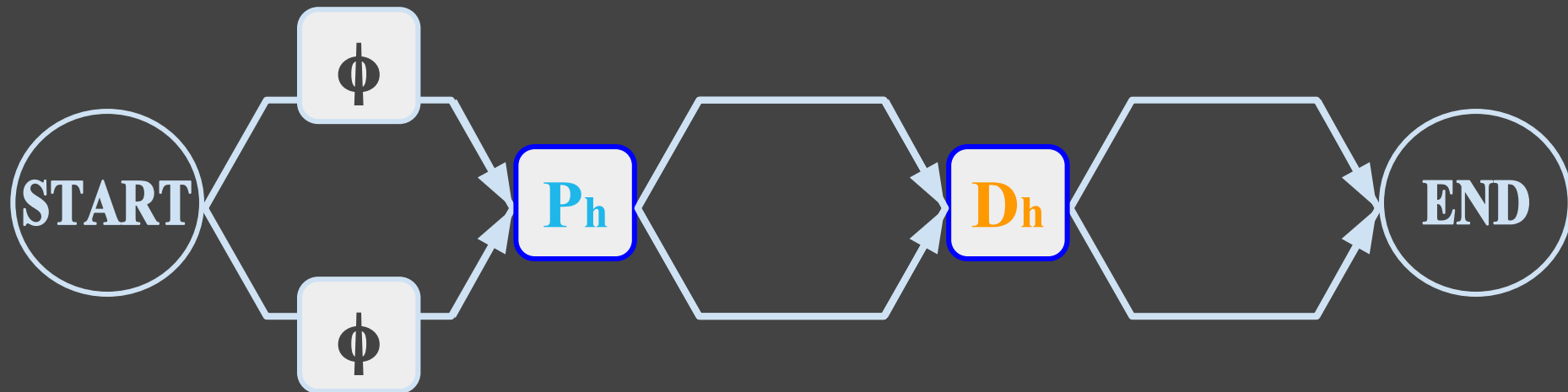
P_h P_s **D_h** D_s

P_h **D_h** P_r D_r



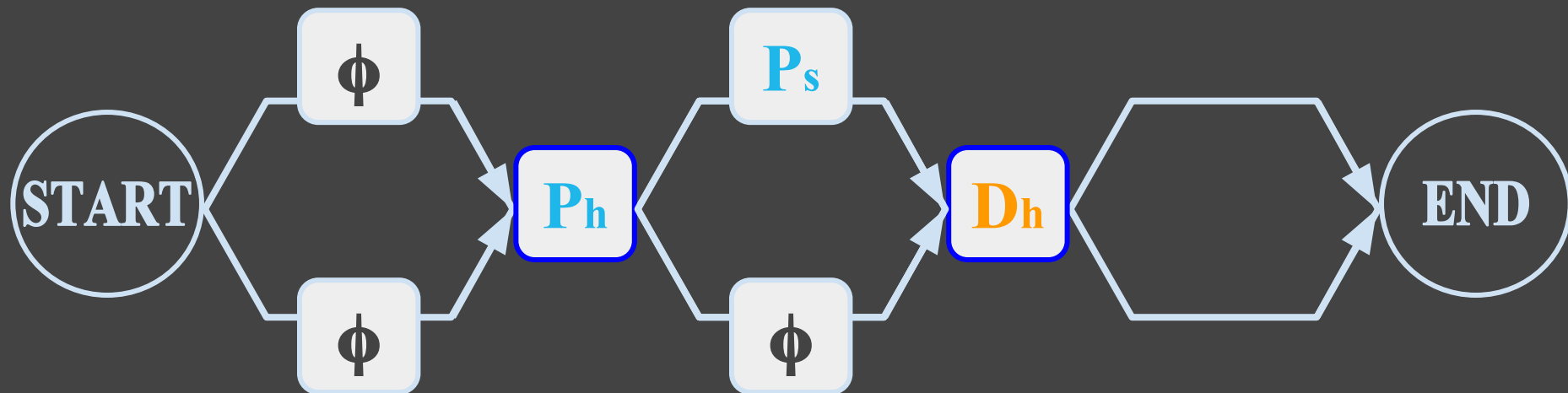
P_h P_s D_h D_s

P_h D_h P_r D_r



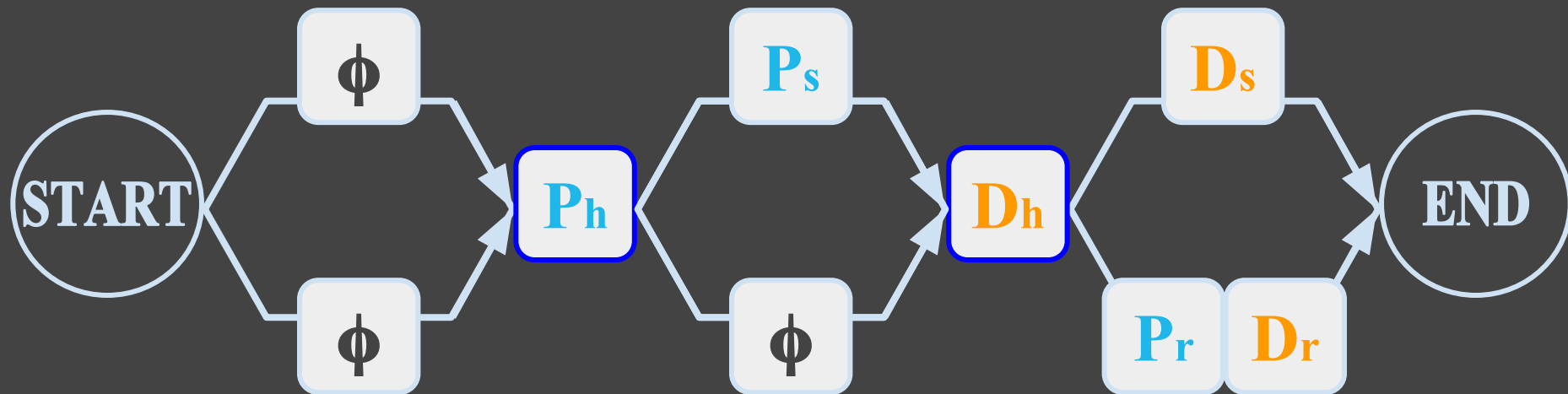
P_h P_s D_h D_s

P_h D_h P_r D_r



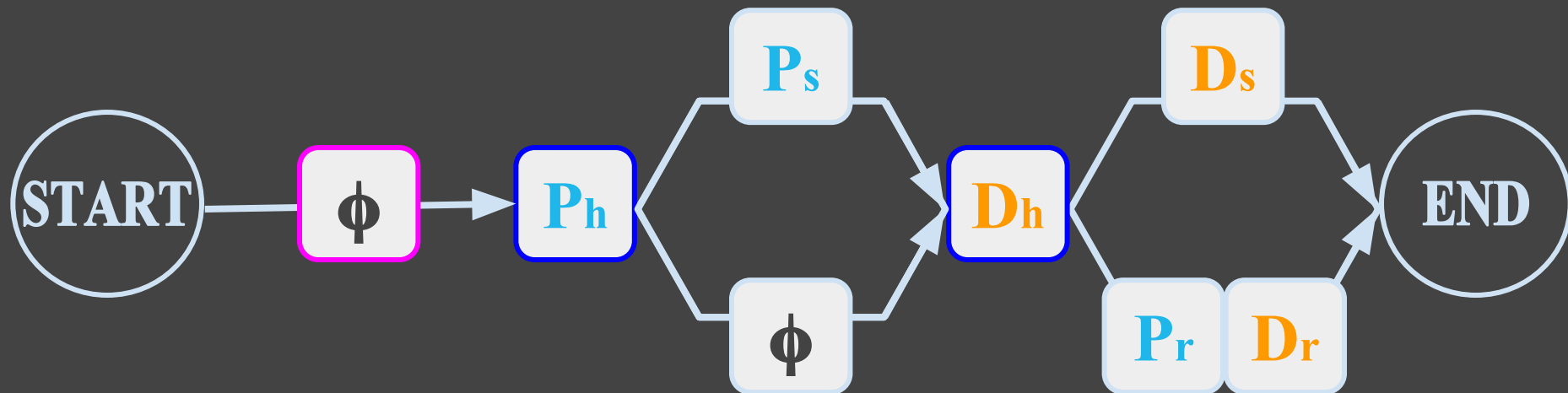
P_h P_s D_h D_s

P_h D_h P_r D_r



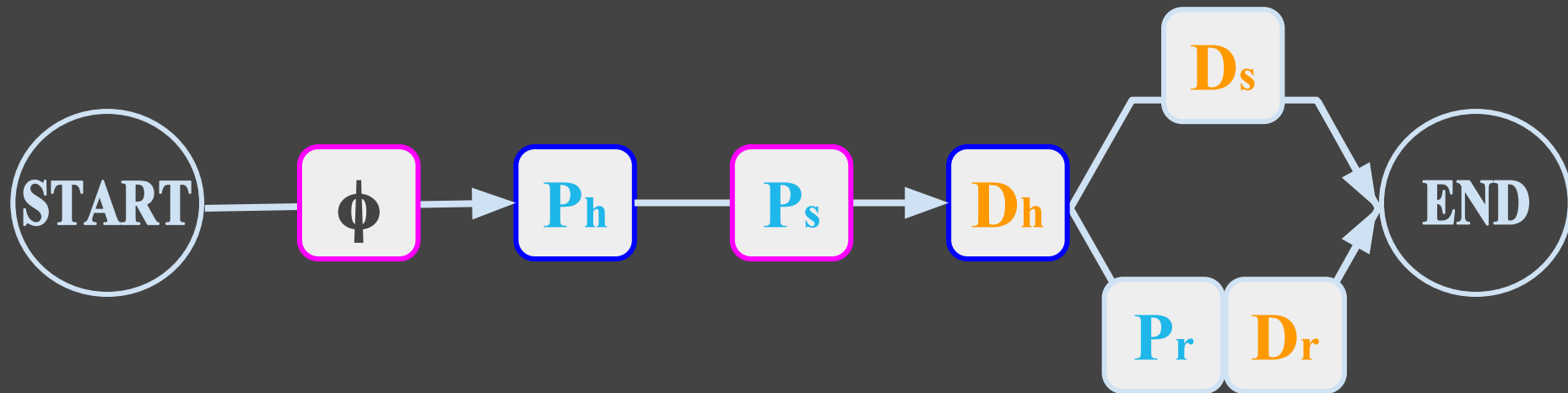
P_h P_s D_h D_s

P_h D_h P_r D_r



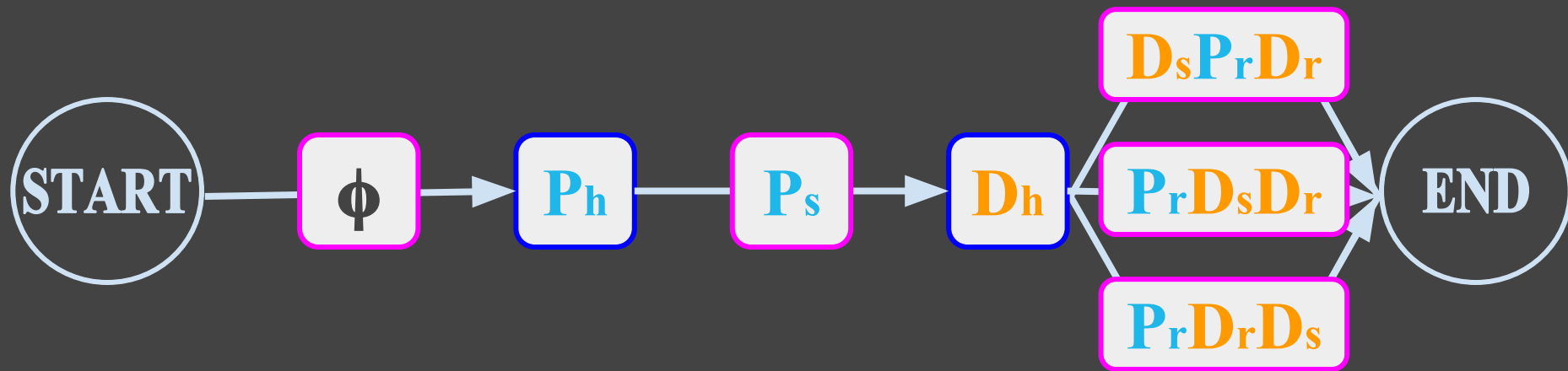
P_h P_s D_h D_s

P_h D_h P_r D_r



$P_h P_s D_h D_s$

$P_h D_h P_r D_r$



Ph Ps Dh Ds

Ph Dh Pr Dr

Sequence alignment result:

Ph Ps Dh Ds Pr Dr

Ph Ps Dh Pr Ds Dr

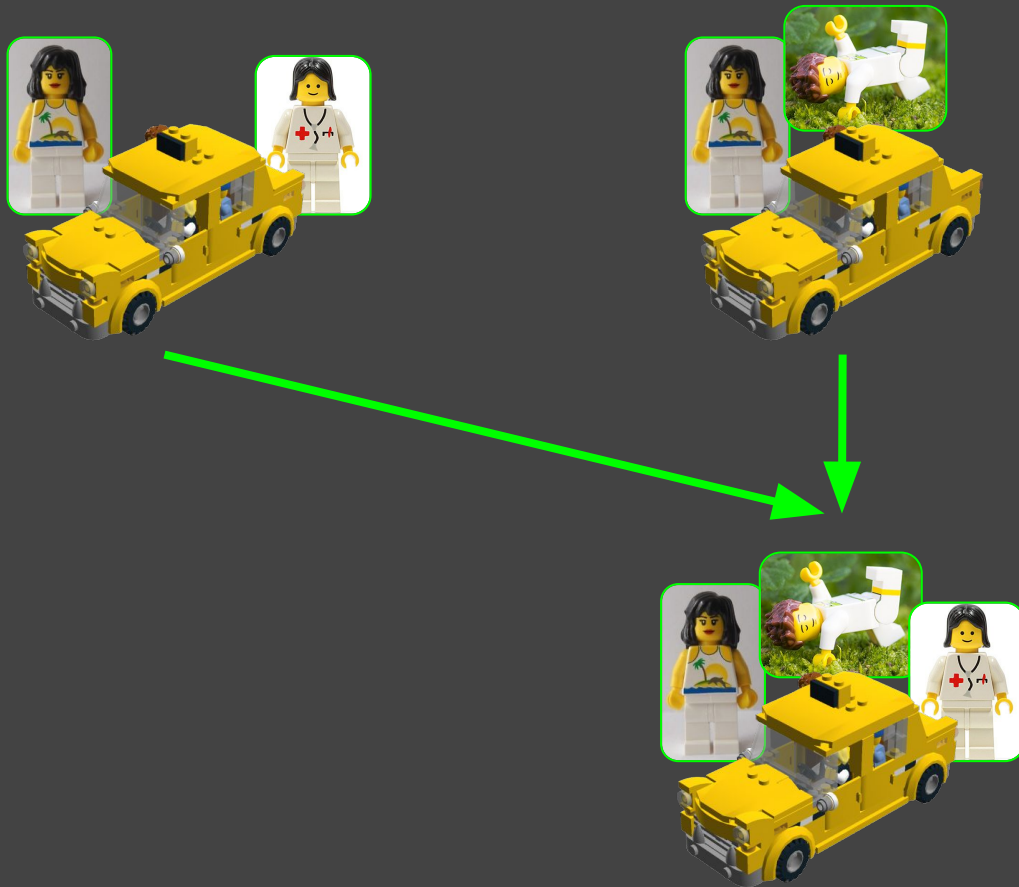
Ph Ps Dh Pr Dr Ds

לחתוך את עלויות המחשוב פי 1000

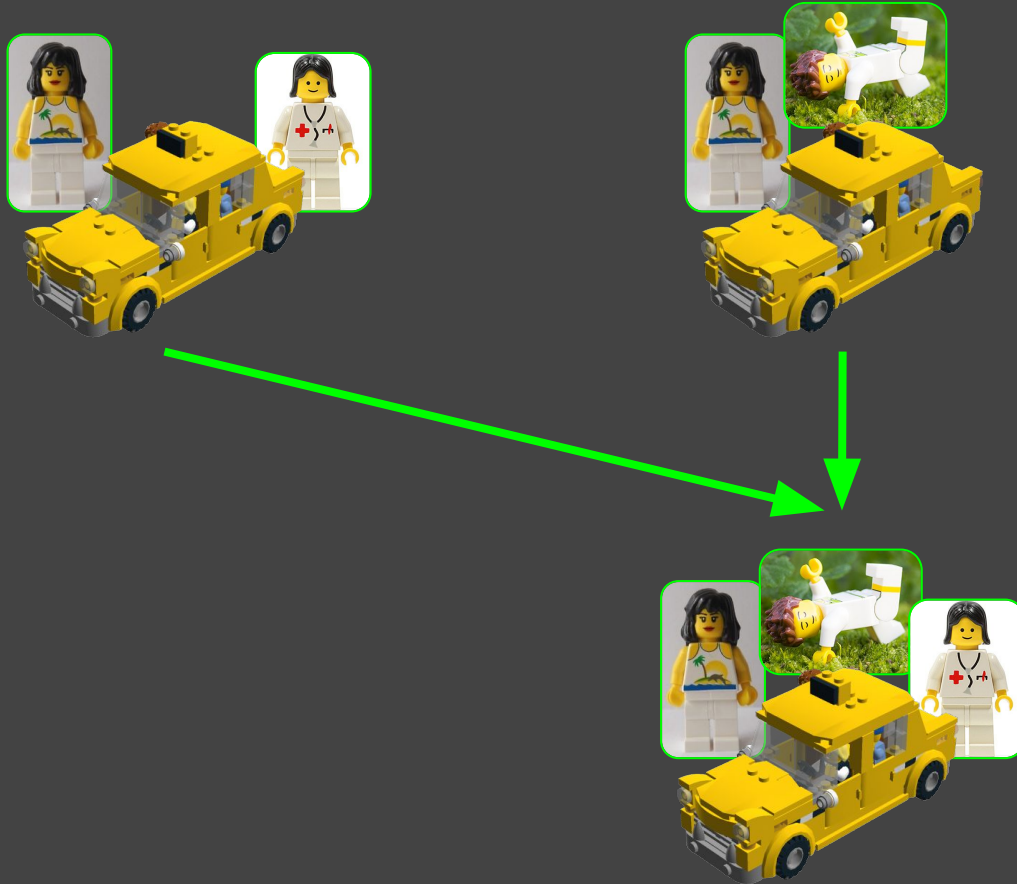


קולולו קולולו הכסף מרשרש מרשרש

האתר המרכזי של כל דבר קולולו



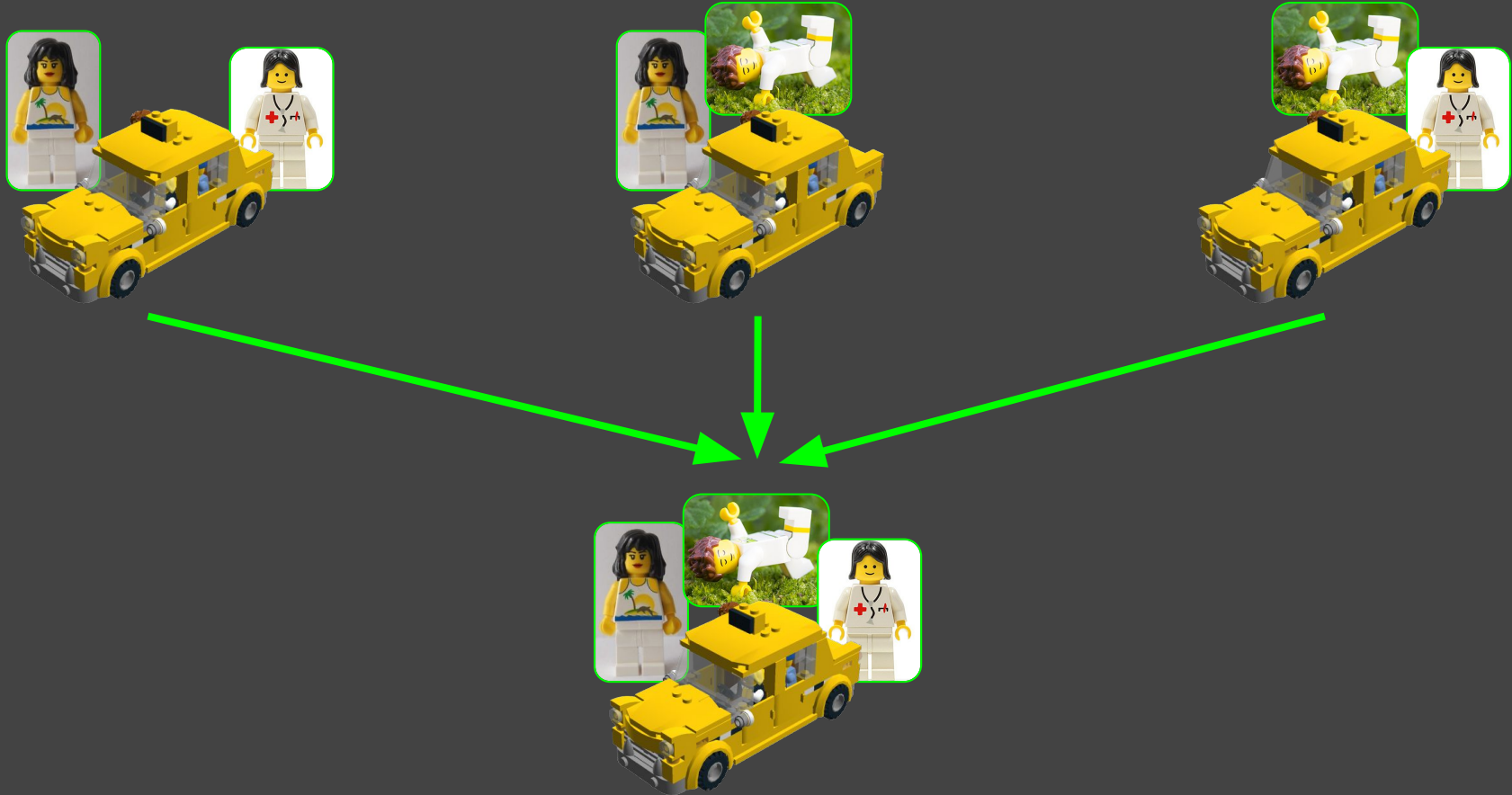
Inductive Ride Sharing



Inductive Ride Sharing



Inductive Ride Sharing



Induction Step

hs

$P_h P_s D_h D_s$

$P_s D_s P_h D_h$

hr

$P_h D_h P_r D_r$

$P_h P_r D_h D_r$

Induction Step

hs

$P_h P_s D_h D_s$

$P_s D_s P_h D_h$

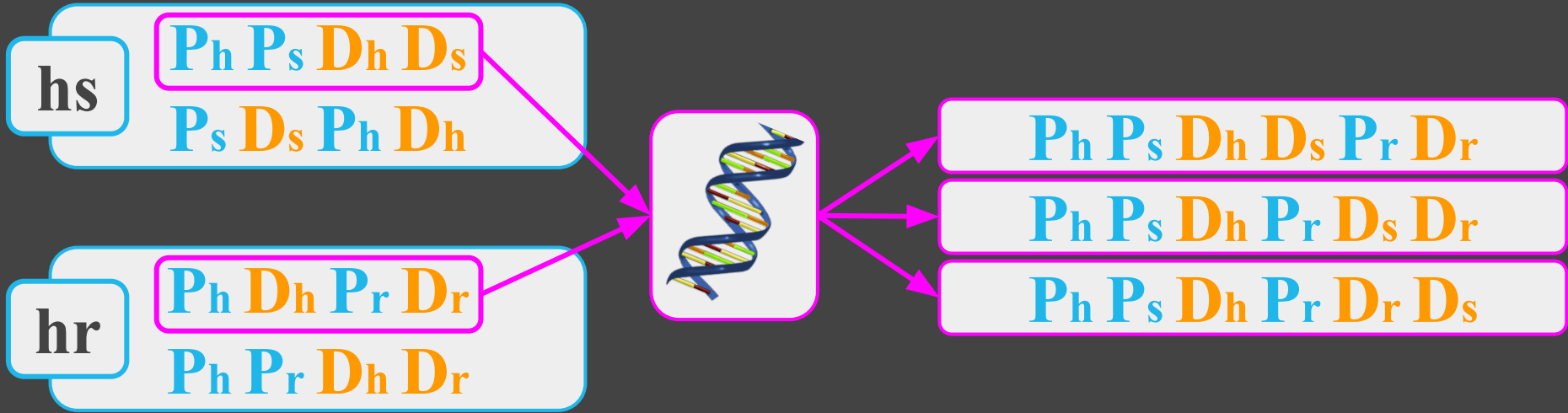
hr

$P_h D_h P_r D_r$

$P_h P_r D_h D_r$



Induction Step



Induction Step

hs

$P_h P_s D_h D_s$
 $P_s D_s P_h D_h$

hr

$P_h D_h P_r D_r$
 $P_h P_r D_h D_r$

sr

$P_s P_r D_s D_r$
 $P_s D_s P_r D_r$



Induction Step

hs

$P_h P_s D_h D_s$
 $P_s D_s P_h D_h$

hr

$P_h D_h P_r D_r$
 $P_h P_r D_h D_r$

sr

$P_s P_r D_s D_r$
 $P_s D_s P_r D_r$



$P_h P_s D_h D_s P_r D_r$

$P_h P_s D_h P_r D_s D_r$

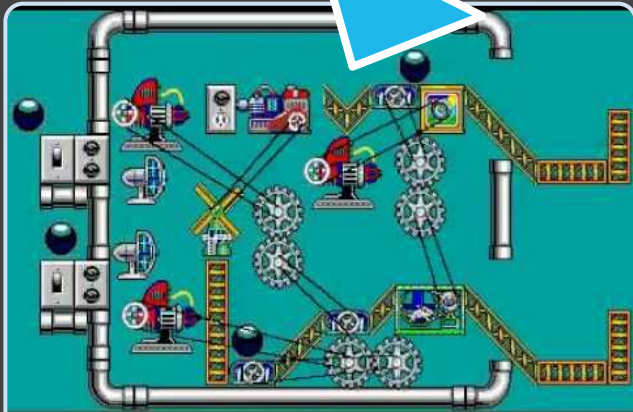
$P_h P_s P_r D_h D_s D_r$

$P_h P_r P_s D_h D_s D_r$

$P_s D_s P_h D_h P_r D_r$

$P_s D_s P_h P_r D_h D_r$

Induction Step



The Ride Sharing Machine

$P_h P_s D_h D_s P_r D_r$

$P_h P_s D_h P_r D_s D_r$

$P_h P_s P_r D_h D_s D_r$

$P_h P_r P_s D_h D_s D_r$

$P_s D_s P_h D_h P_r D_r$

$P_s D_s P_h P_r D_h D_r$

Induction Step

6 OUT OF 90 ?!?



Dr

Dr

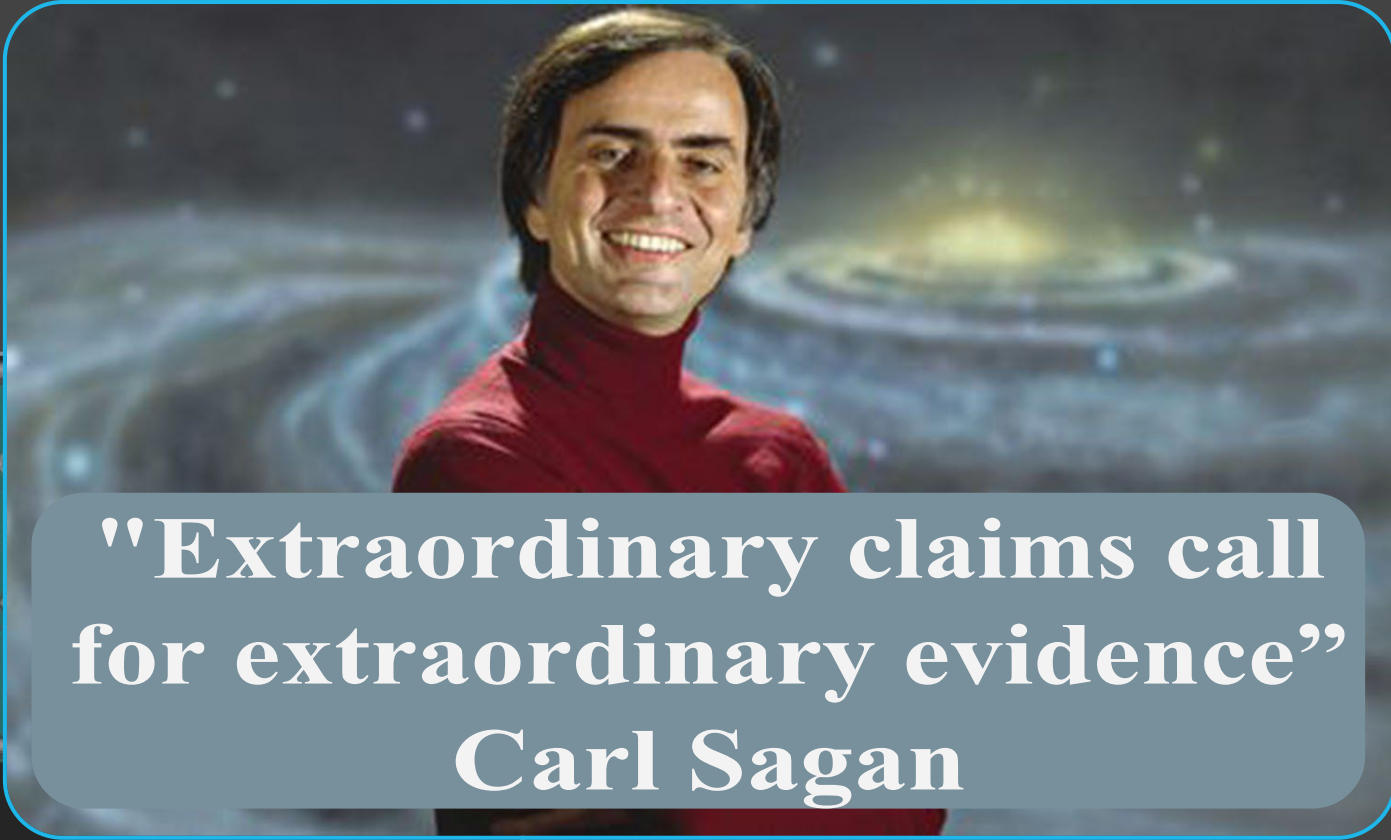
Dr

Dr

Dr

Dr

Induction Step



**"Extraordinary claims call
for extraordinary evidence"**
Carl Sagan

r Dr

s Dr

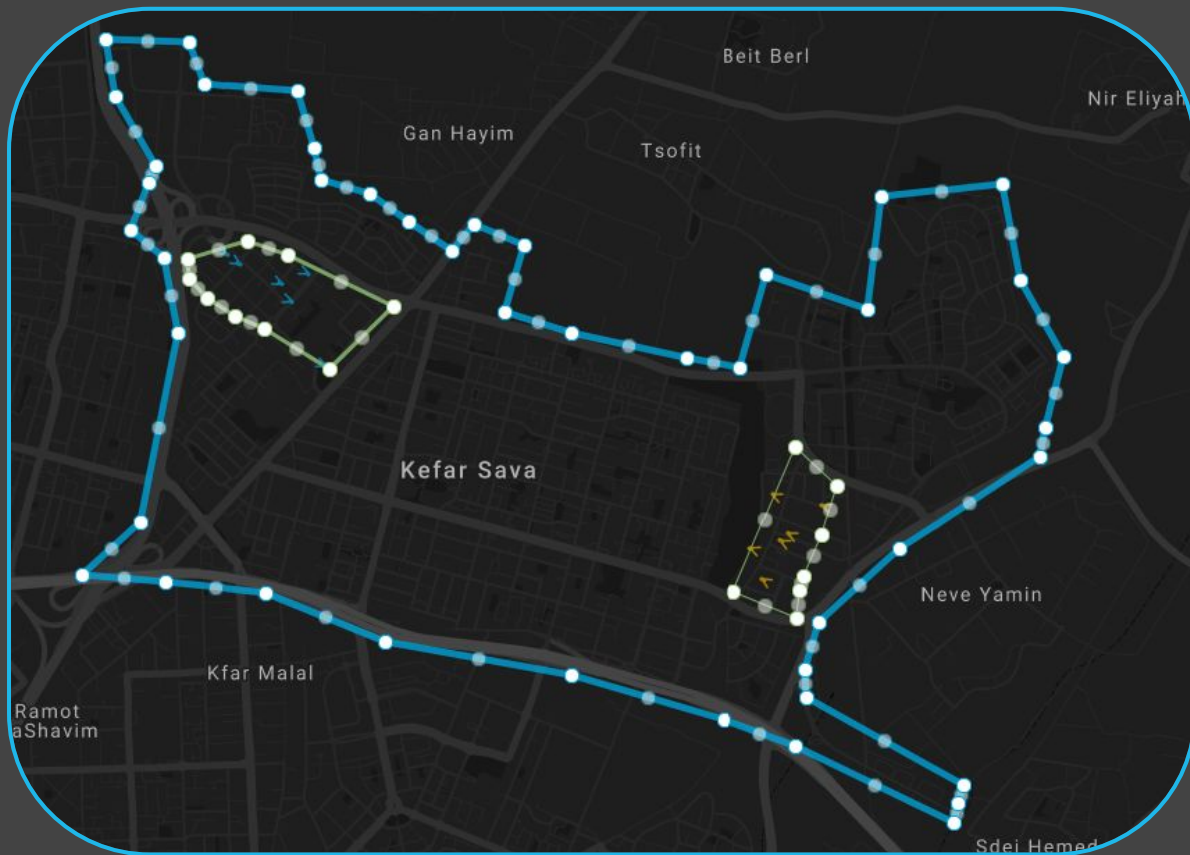
s Dr

s Dr

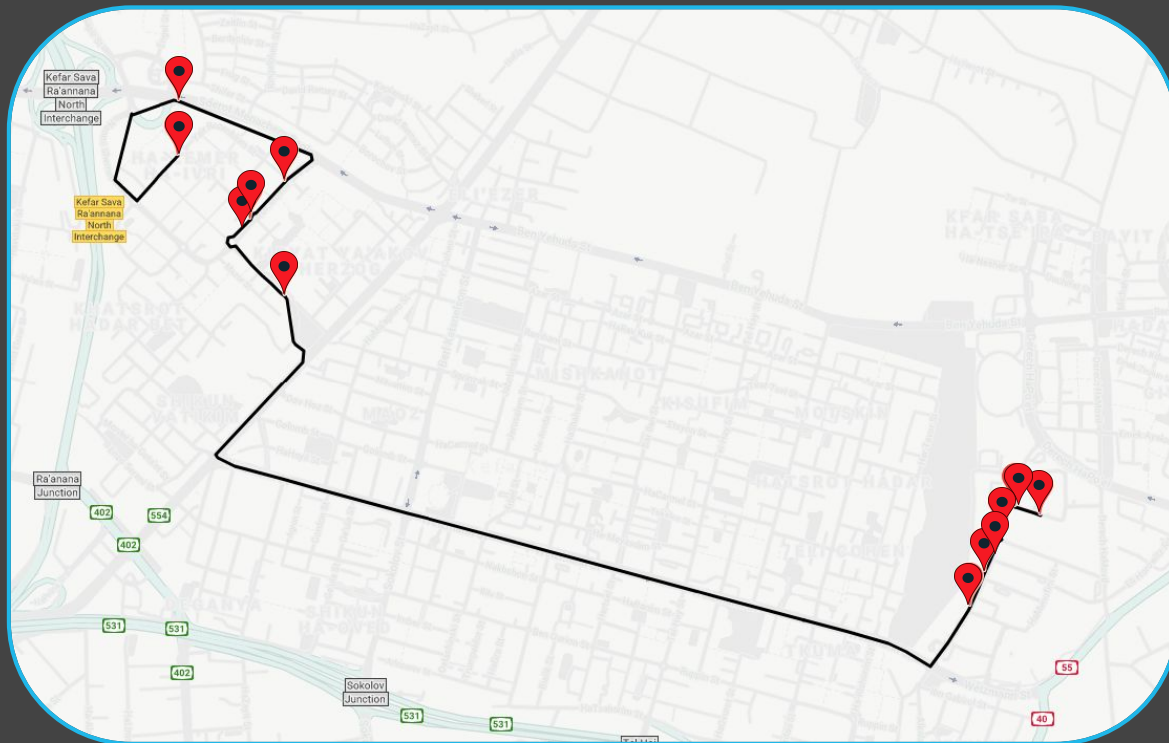
r Dr

h Dr

Results in the Wild



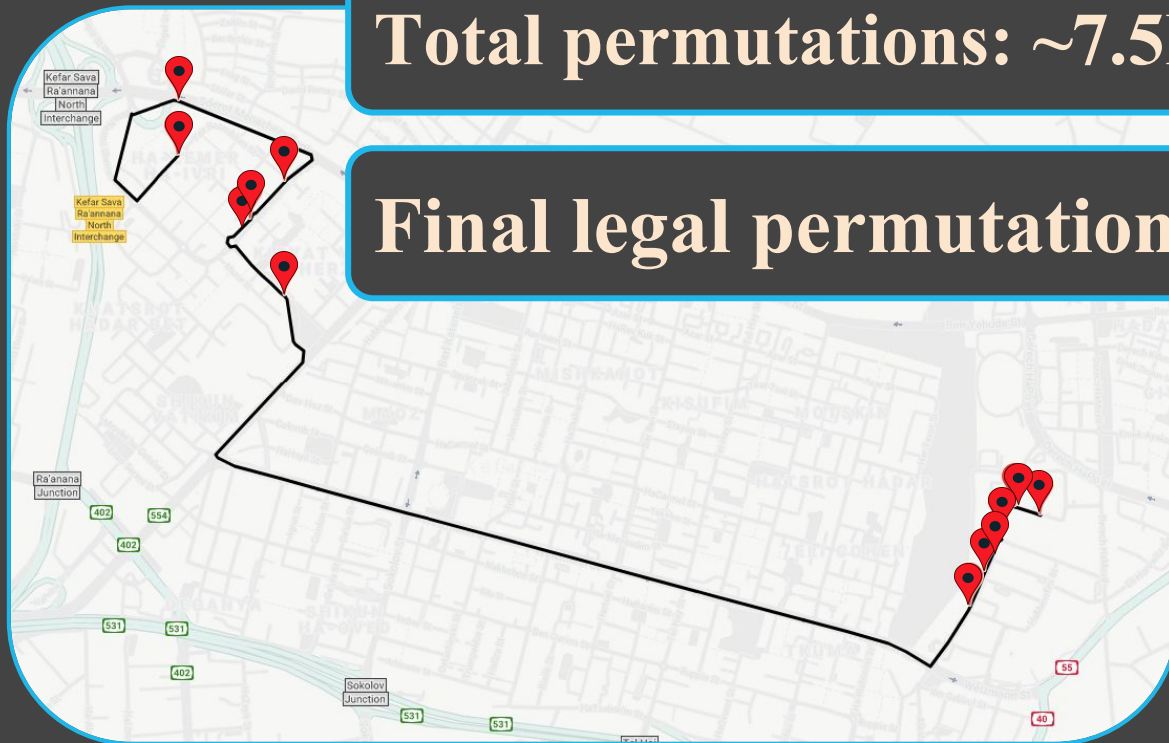
Results in the Wild



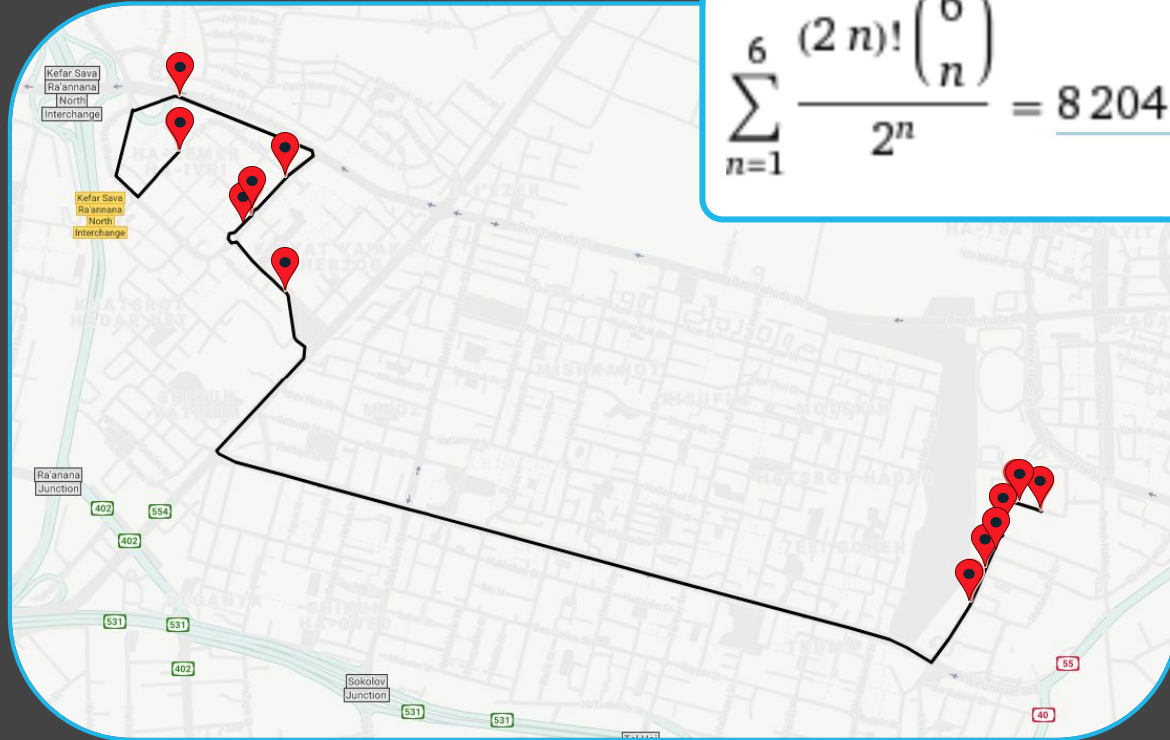
Results in the Wild

Total permutations: ~7.5M

Final legal permutations: 240



Results in the Wild

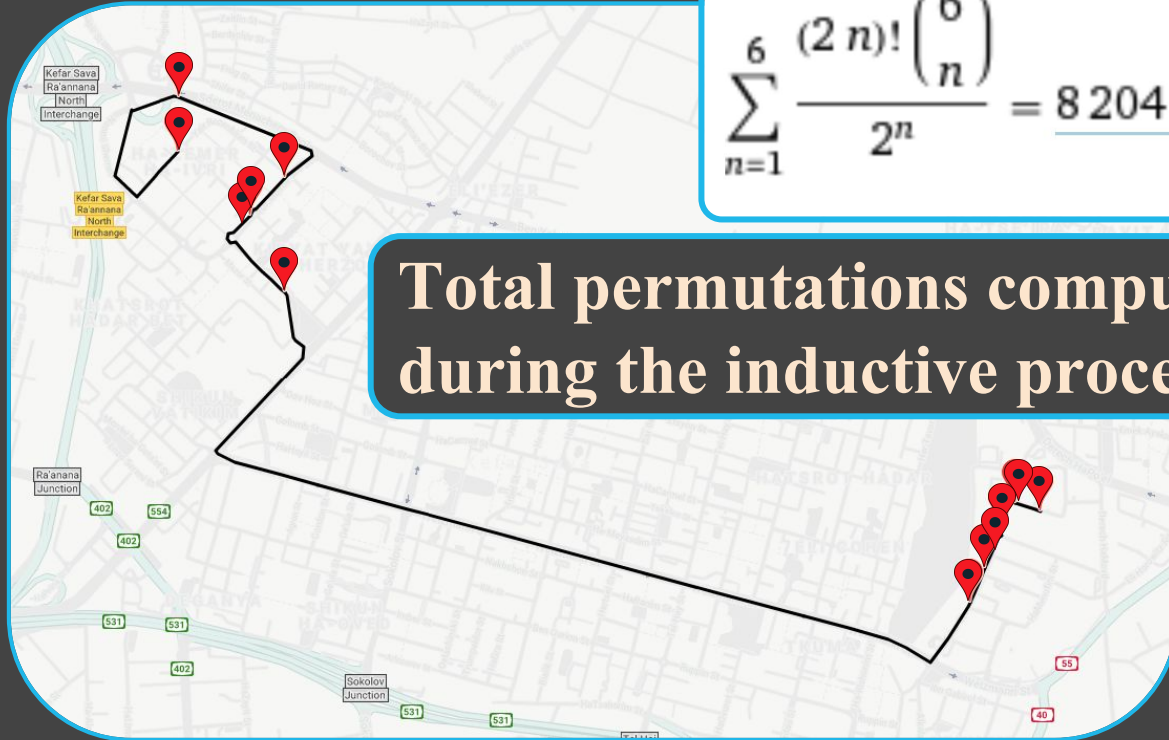


$$\sum_{n=1}^6 \frac{(2n)! \binom{6}{n}}{2^n} = \underline{8\,204\,496}$$

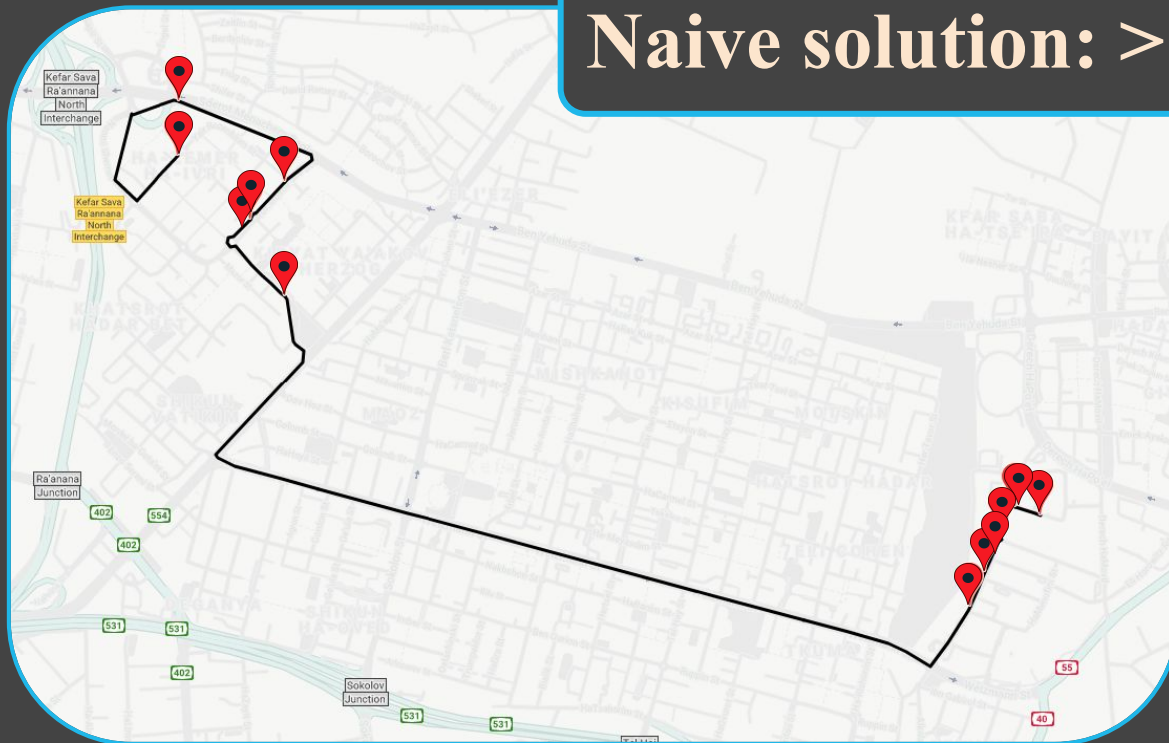
Results in the Wild

$$\sum_{n=1}^6 \frac{(2n)! \binom{6}{n}}{2^n} = \underline{8\,204\,496}$$

Total permutations computed during the inductive process: 3822



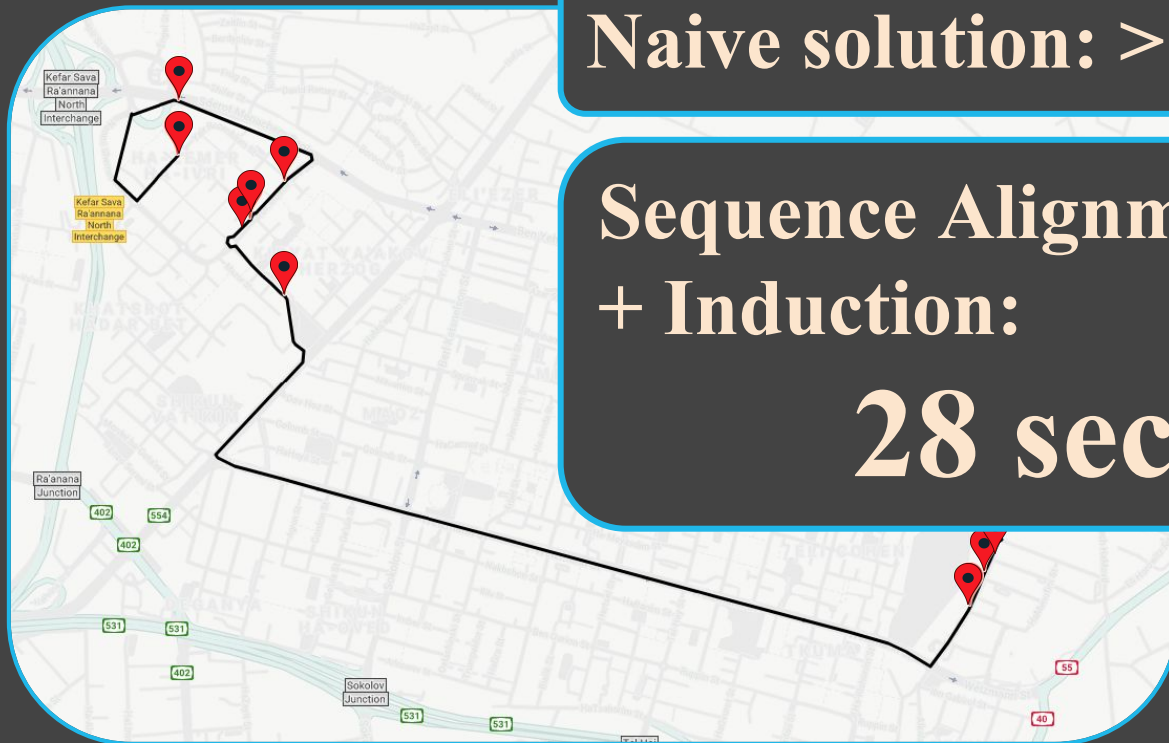
Naive solution: > 1 hour



Results in the Wild

Naive solution: > 1 hour

Sequence Alignment
+ Induction:
28 sec.



Take Home Message



Thank you :)

Questions?

DalyaG@gmail.com

