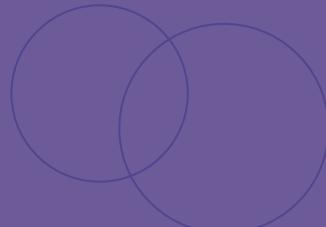
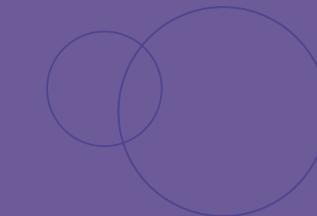


# Music Transformer

## Generating music with long-term structure

Paper: Google Brain



# Table of contents

01 Introduction

02 Related Work

03 Model

04 Experiments





01

# Jntroduction



# Goal

Generating Music using Relative Self Attention  
Transformer to Capture Relative in Music



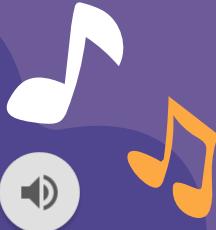
# Dataset

- J.S. Bach Chorales
- Piano-E-Competition



# Challenge!!

- Multiple dimensions
- Motifs
- Relative differences  
arguably : Timing & Pitch



Language

text



speech

Music

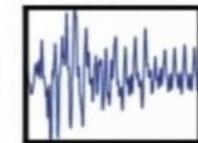
score



performance



sound



composer

performer

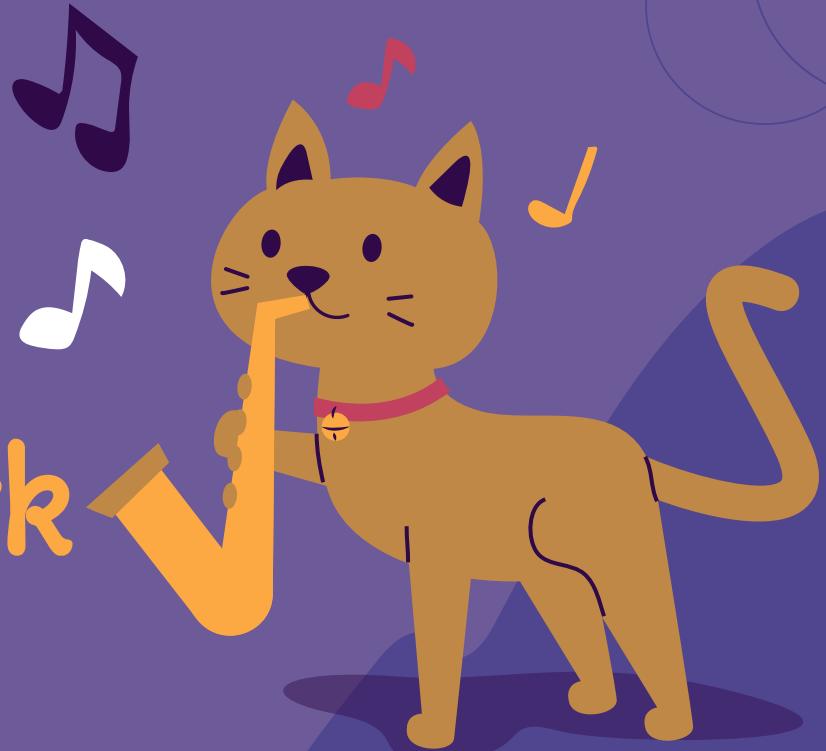
instrument

listener

(Image from Simon & Oore, 2016)

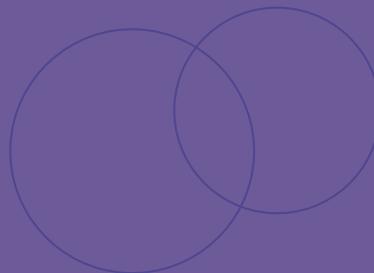
02

# Related Work



# 2002

Hidden Markov model



# 2016

RNNs

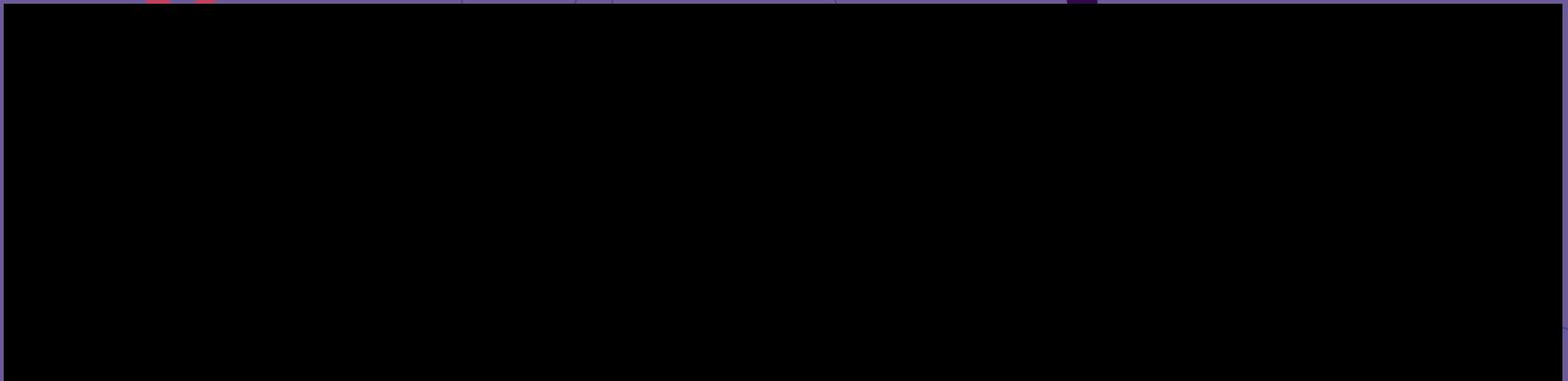


# 2017

Bidirectional LSTMs



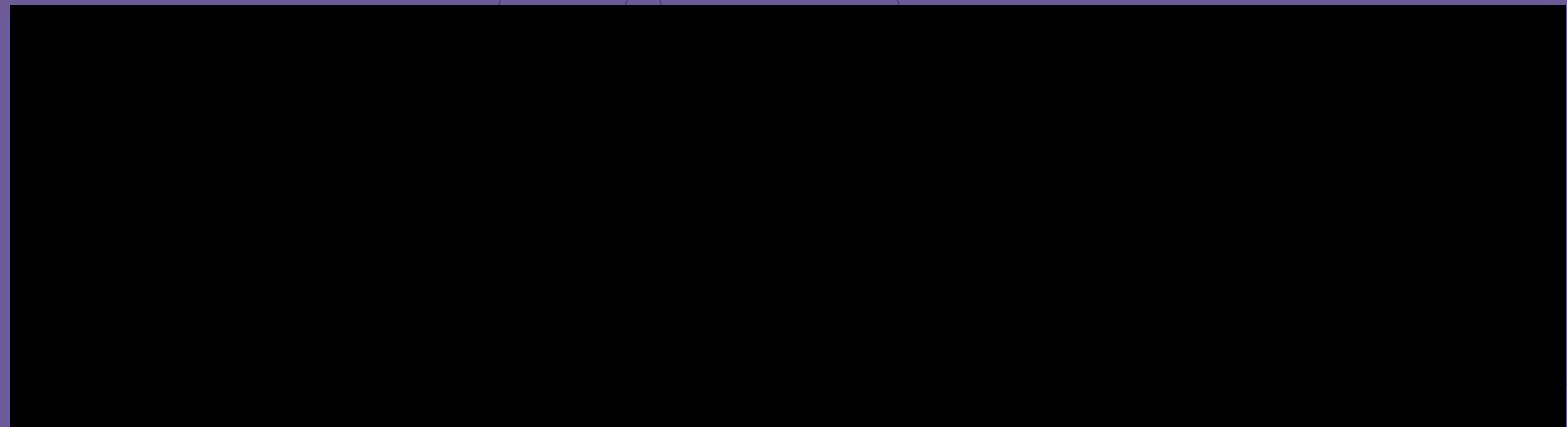
# Performance RNN



# 2018

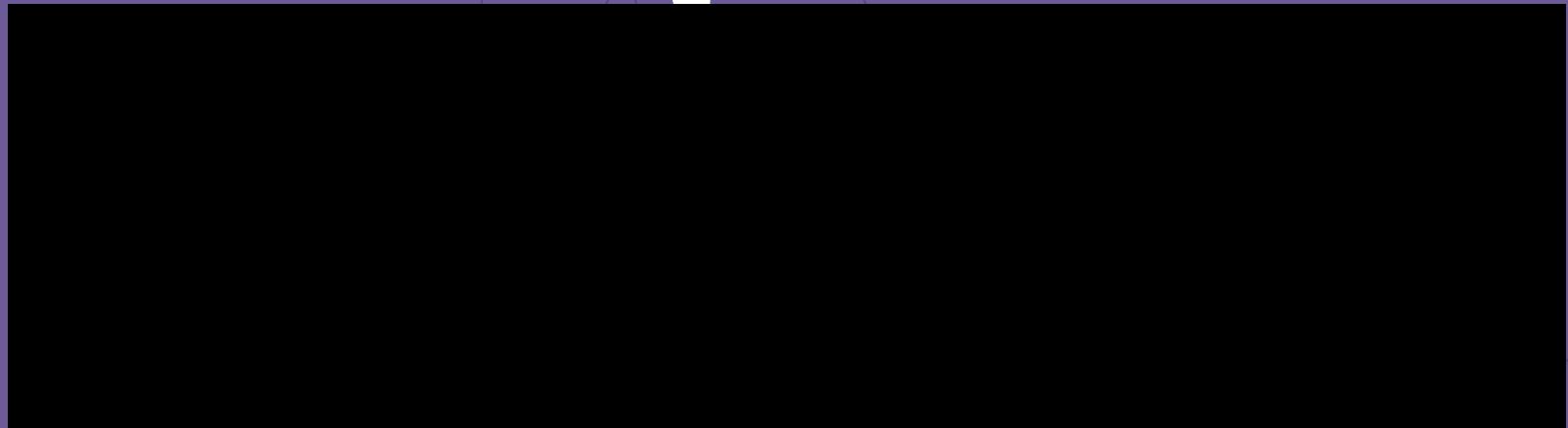
Self-attention in Transformer

# Vanilla Transformer



Source: <https://magenta.tensorflow.org/music-transformer>

# Music Transformer



Given  
motif



RNN-LSTM



Transformer



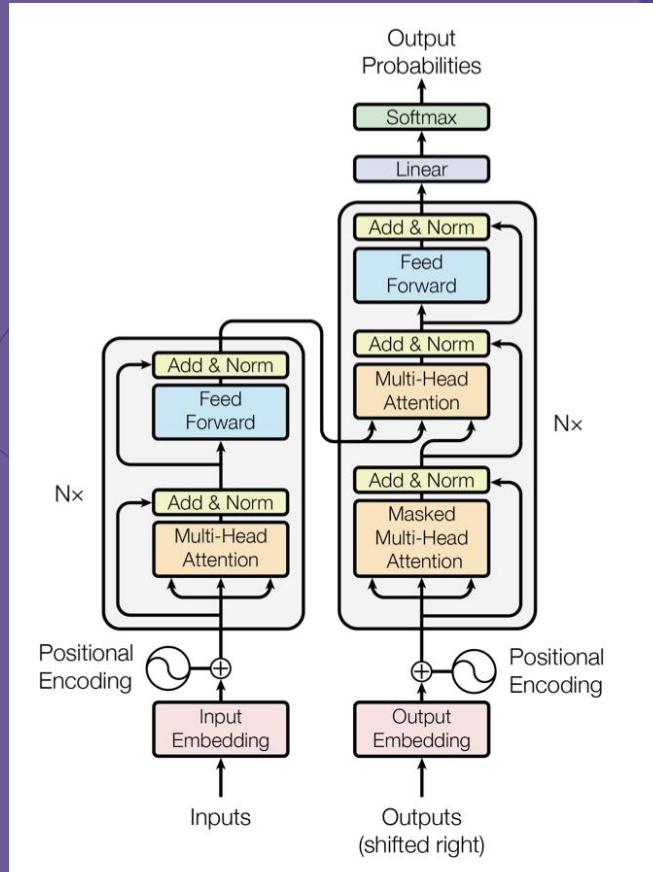
Music  
Transformer



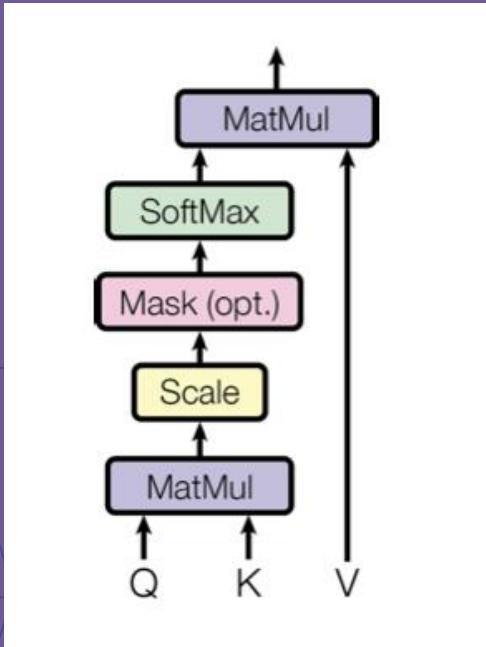


# 03 Model

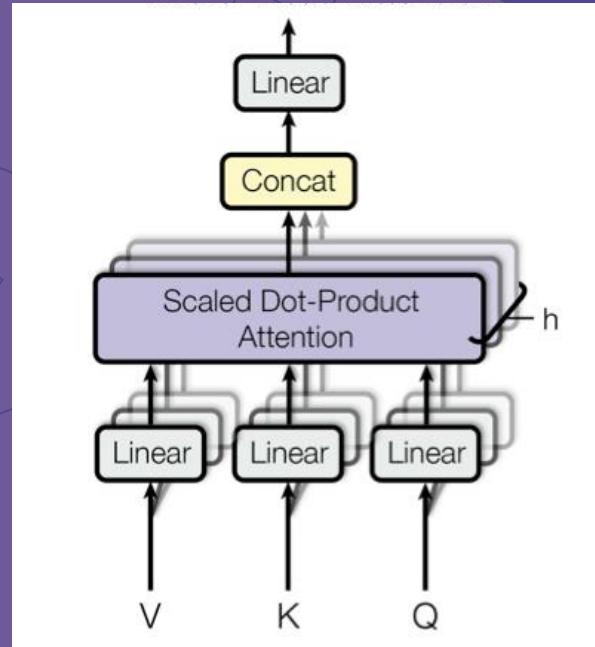
# The Transformer Model architecture



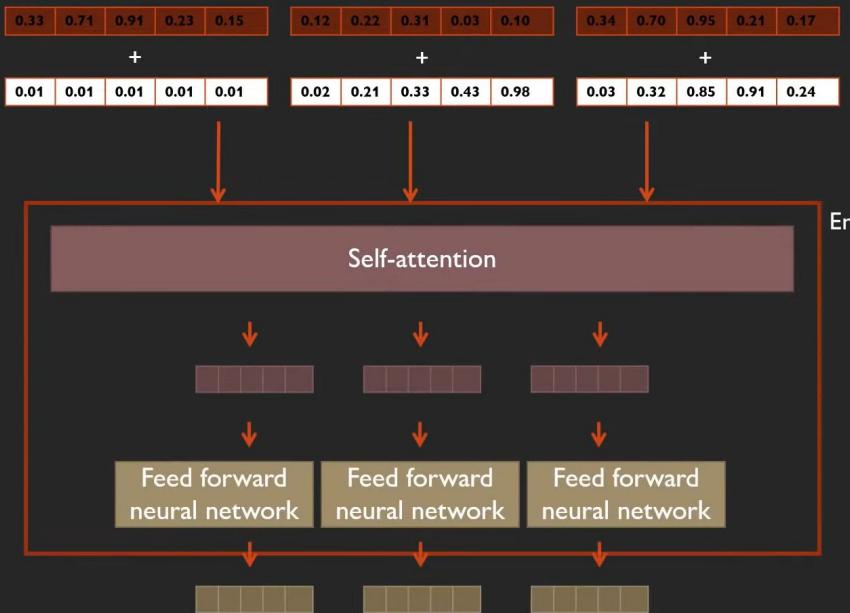
# Scaled Dot-Product Attention



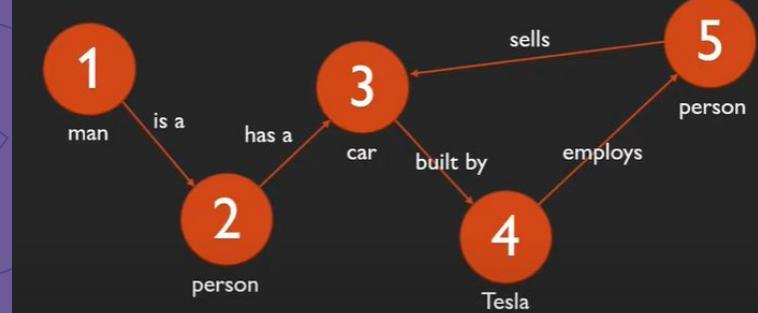
# Multi-Head Attention

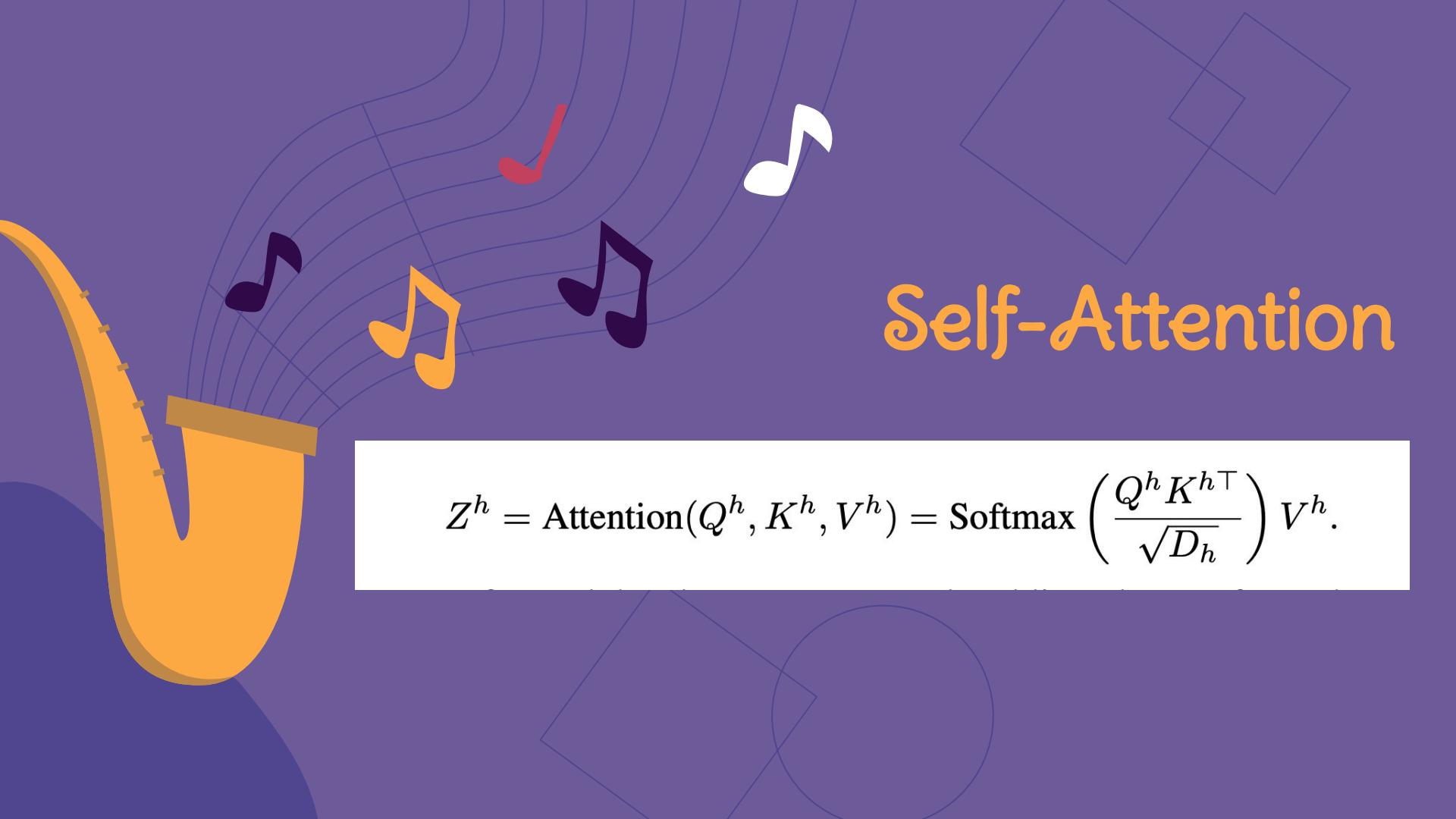


# Queen and king



Encoder





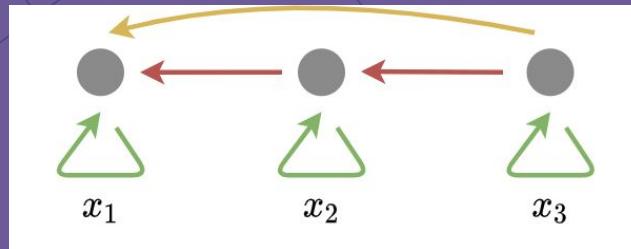
# Self-Attention

$$Z^h = \text{Attention}(Q^h, K^h, V^h) = \text{Softmax} \left( \frac{Q^h K^{h\top}}{\sqrt{D_h}} \right) V^h.$$

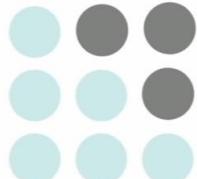
# Relative Position Self-Attention

$$Z^h = \text{Attention}(Q^h, K^h, V^h) = \text{Softmax} \left( \frac{QK^\top + S^{rel}}{\sqrt{D_h}} \right) V.$$

# Relative Position Self-Attention



$$\text{softmax}(QK^\top + Qf(E_{rel}))$$



Modulated by  
relative positions



$$Qf(E_{rel})$$

Modulated by  
relative positions

0,0		
1,0	1,1	
2,0	2,1	2,2

$$Qf(E_{rel})$$

Modulated by  
relative positions

0		
-1	0	
-2	-1	0

# Previous work $O(L^2D)$ : **8.5 GB** per layer (Shaw et al, 2018)

Per layer,  $L=2048$ ,  $D=512$

$$\text{softmax}(QK^\top + Qf(E_{rel}))$$

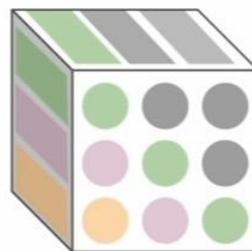
Relative  
embeddings  $E_{rel}$



Relative distances

0		
-1	0	
-2	-1	0

$f(E_{rel})$



R

Multiply by Q



Our formulation O(LD): **4.2 MB** per layer

$$\text{softmax}(QK^\top + \text{skew}(QE_{\text{rel}}^\top))$$

Per layer, L=2048, D=512

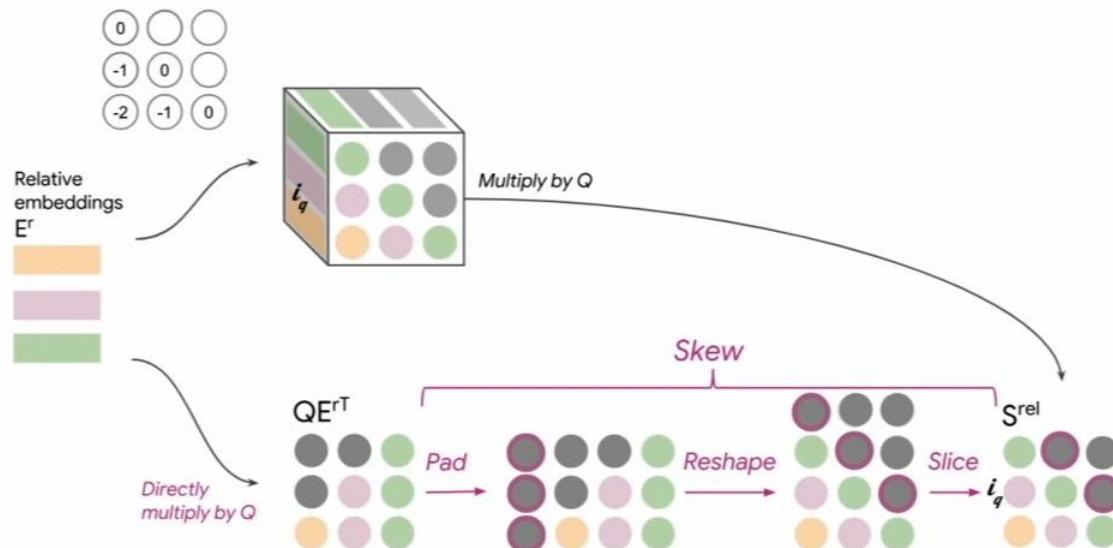


# Skewing to reduce relative memory from $O(L^2D)$ to $O(LD)$

Per layer,  $L=2048$ ,  $D=512$

Previous work  
 $O(L^2D): 8.5 \text{ GB}$

Our work  
 $O(LD): 4.2 \text{ MB}$





# Relative Local Attention

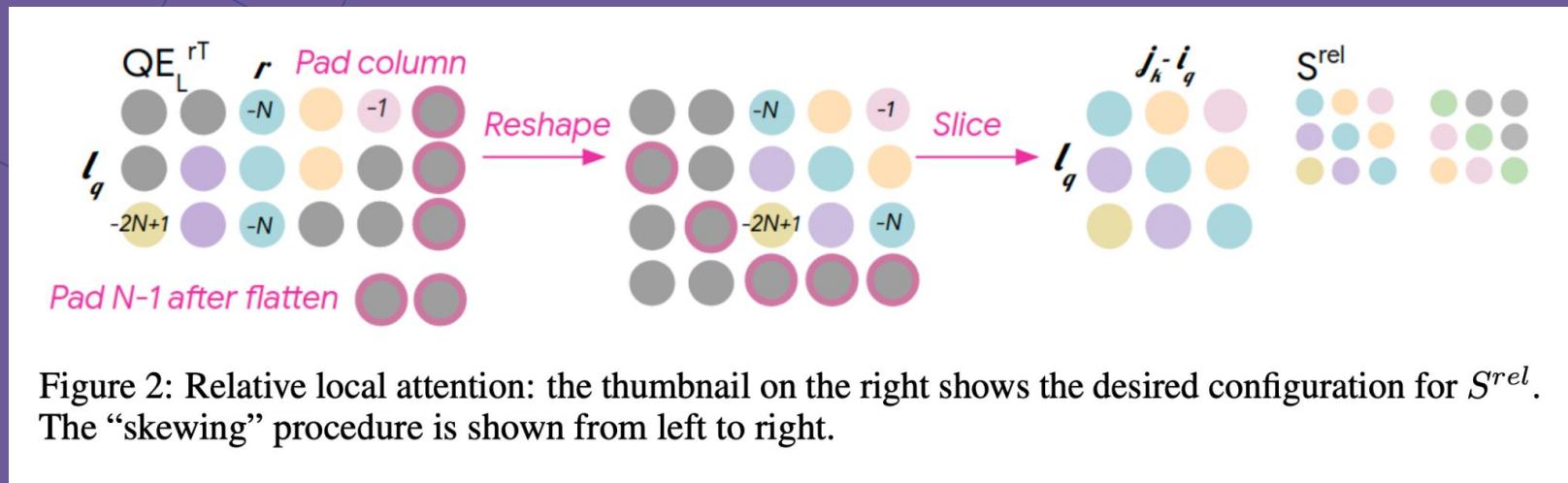
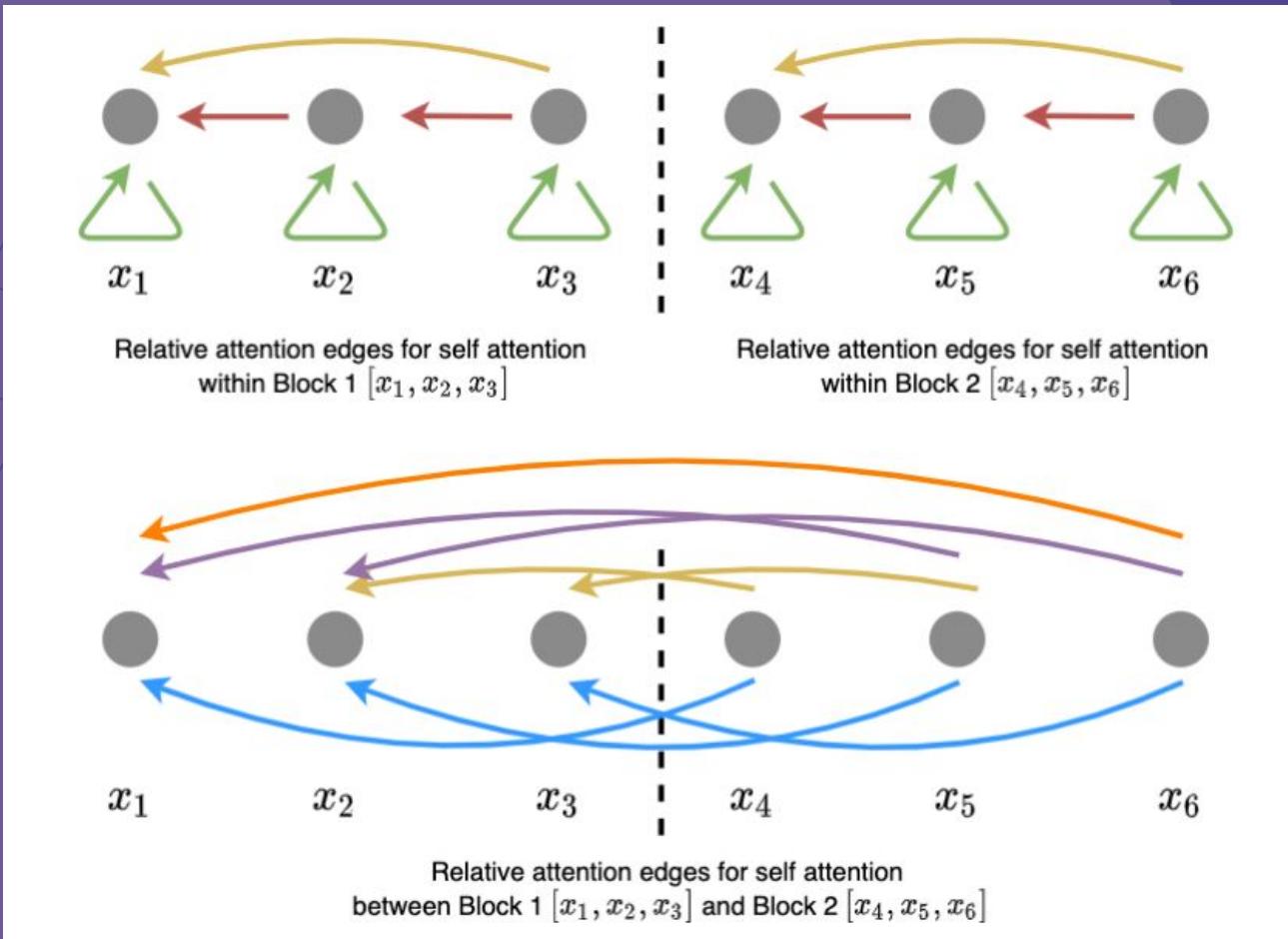
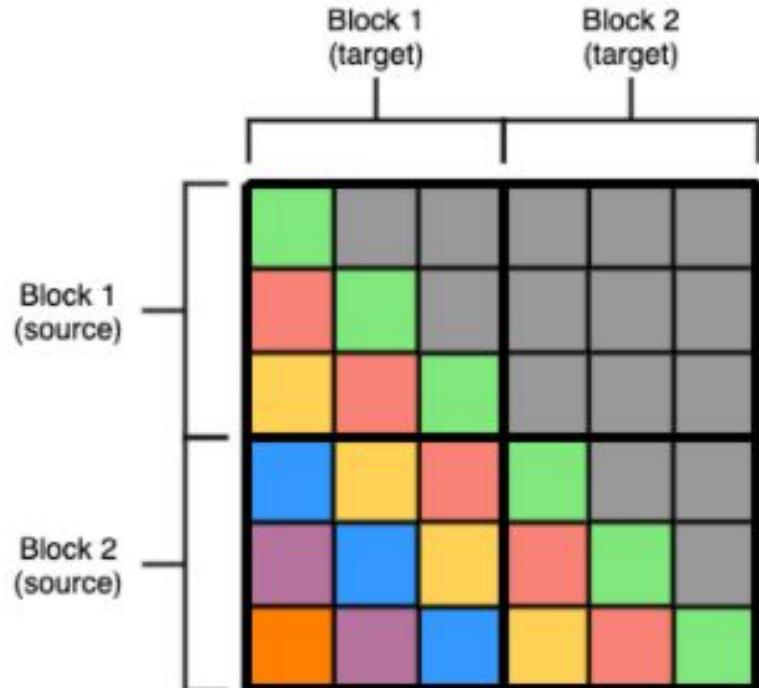


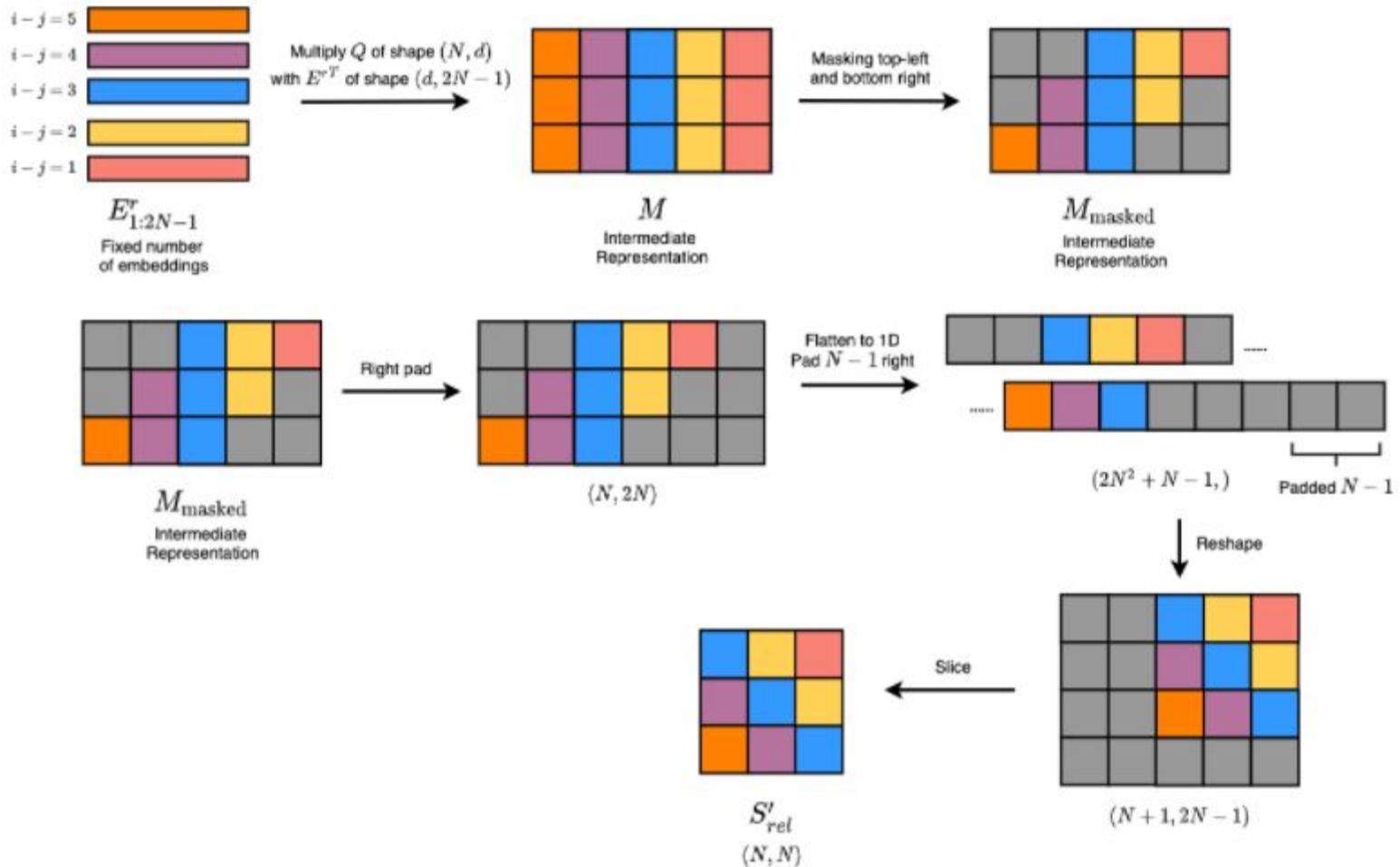
Figure 2: Relative local attention: the thumbnail on the right shows the desired configuration for  $S^{rel}$ . The “skewing” procedure is shown from left to right.





$$S_{rel}$$

Matrix with relative  
position representations

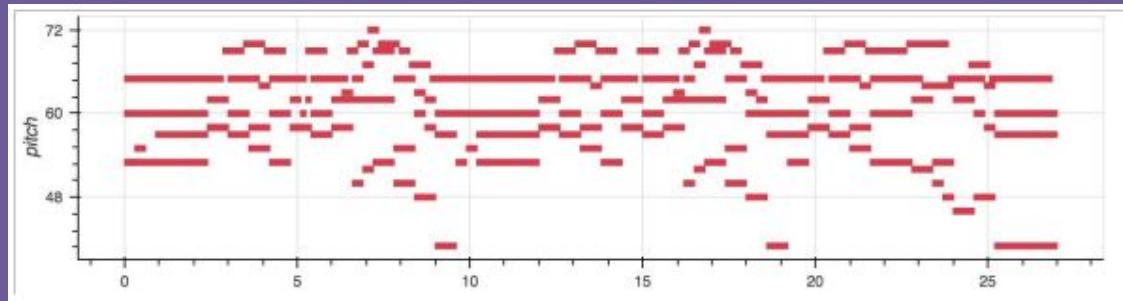


04

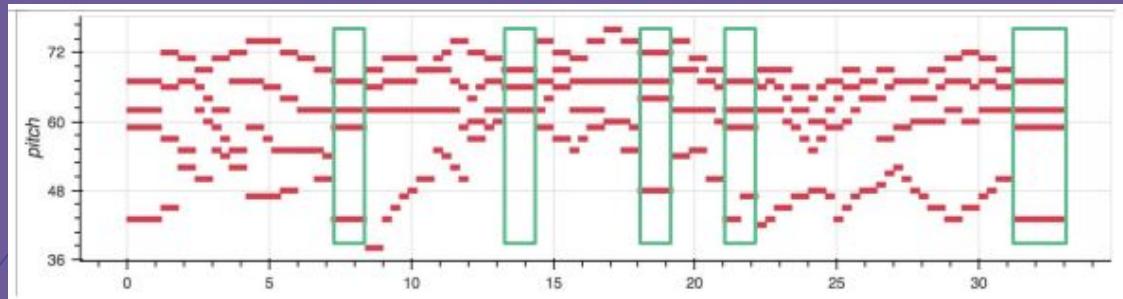
# Evaluation



# J.S. BACH Chorales



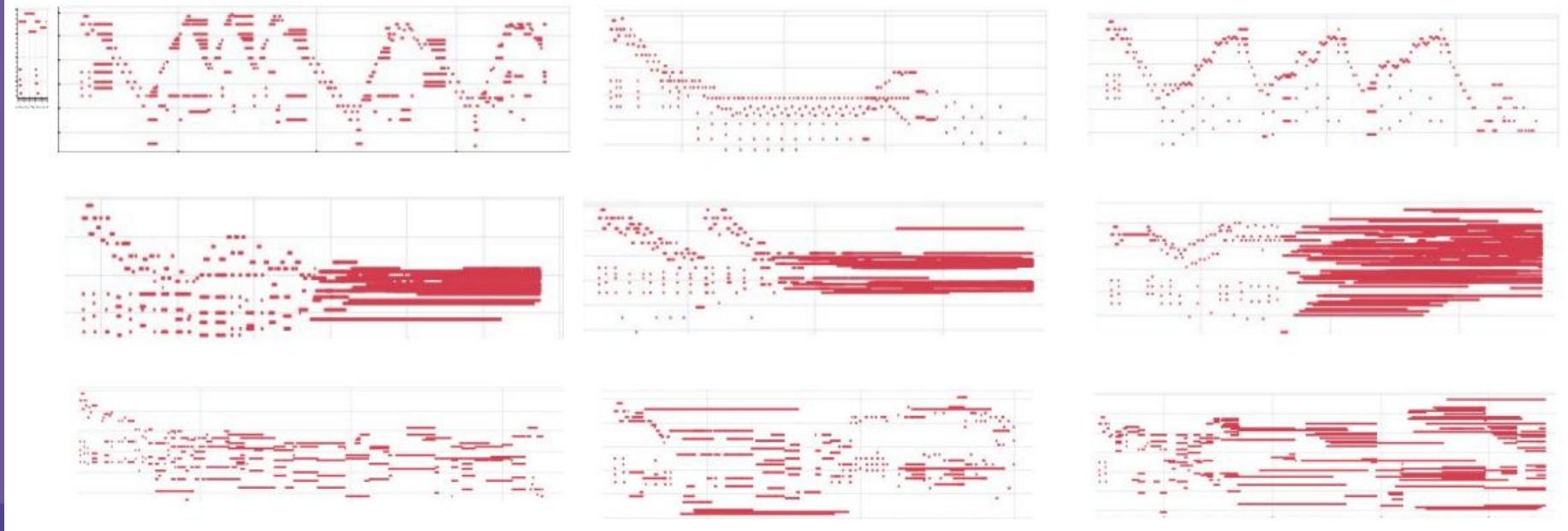
Transformer **without** relative self-attention



Transformer **with** relative self-attention.

Green vertical boxes indicate the endings of (sub)phrases where cadences are held.

# PJANO-E-Competition



Comparing how models continue a prime (top left). Repeated motives and structure are seen in samples from Transformer with relative attention (top row), but less so from baseline Transformer (middle row) and PerformanceRNN (LSTM) (bottom row).

# Validation NLL

Note-wise validation NLL on J.S.Bach Chorales at 16th notes. Relative attention, more timing and relational information improve performance.

Model variation	Validation NLL
COCONET (CNN, chronological, 64L, 128 3x3f)	0.436
COCONET (CNN, orderless, 64L, 128 3x3f)	$\leq 0.238^6$
Transformer (TF) baseline (Vaswani et al., 2017) (5L, 256hs, 256att, 1024ff, 8h)	0.417
TF baseline + concat positional sinusoids (cps)	0.398
TF baseline + concat positional sinusoids, instrument labels (cpsi)	0.370
Relative Transformer (Shaw et al., 2018) (5L, 512hs, 512att, 512ff, 256r, 8h)	0.357
Relative Transformer + concat positional sinusoids, instrument labels (cpsi)	0.347
Relative Transformer + cpsi + relative pitch and time	0.335

Validation NLL for Piano-e-Competition dataset, with event-based representation with lengths L = 2048. Transformer with relative attention (with our efficient formulation) achieves state-of-the-art performance.

Model variation	Validation NLL
PERFORMANCE RNN (LSTM) (3L, 1024hs)	1.969
LSTM with attention (3L, 1024hs, 1024att)	1.959
Transformer (TF) baseline (6L, 256hs, 512att, 2048fs, 1024r, 8h)	1.861
TF with local attention (Liu et al., 2018) (8L, 1024fs, 512bs)	1.863
TF with relative global attention (our efficient formulation) (6L, 2048fs, 1024r)	<b>1.835</b>
TF with relative local attention (ours) (6L, 1024fs, 2048r, 512bs)	<b>1.840</b>

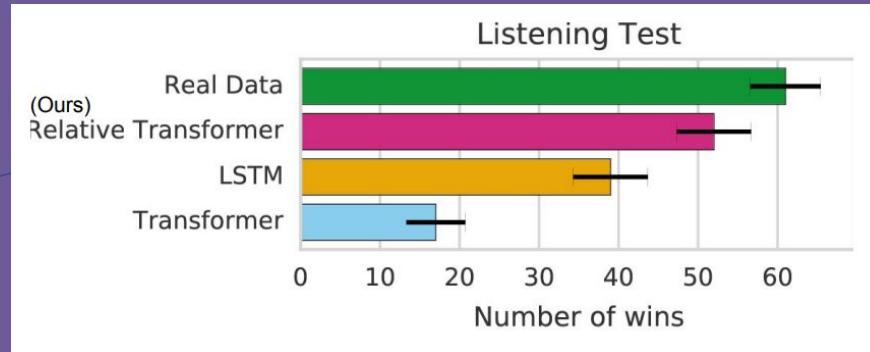
# PJANO-E-Competition

## HARMONIZATION: CONDITIONING ON MELODY

Model variation	NLL
Baseline Transformer	2.066
Relative Transformer (ours)	1.786

Validation conditional NLL given ground truth melody from Piano-e-Competition.

## Human Evaluation



Number of wins for each model. Error bars show standard deviations of mean

# Group Member

- |    |           |                   |            |
|----|-----------|-------------------|------------|
| 1. | Voramate  | Plodprong         | 6310422028 |
| 2. | Supanut   | Thaiprasit        | 6310422034 |
| 3. | Rakchanok | Thongkumpan       | 6310422039 |
| 4. | Varattaya | Rojanarachneekorn | 6310422044 |
| 5. | Piyaboon  | Kunakornjittirak  | 6310422047 |

# Thanks!



# Generalizing Relative Attention To Capture Relational Information

$$S^{rel} = \text{Skew}(QE^r) + Q(R^t + R^p)$$

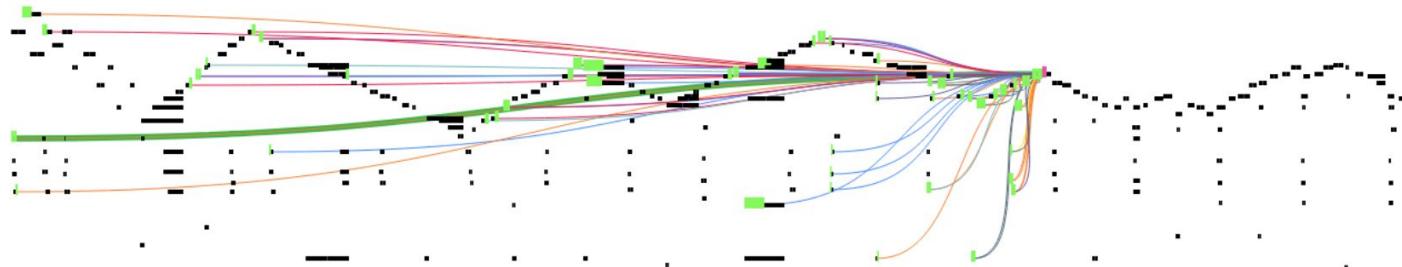


Figure 8: This piece has a recurring triangular contour. The query is at one of the latter peaks and it attends to all of the previous high notes on the peak, all the way to beginning of the piece.

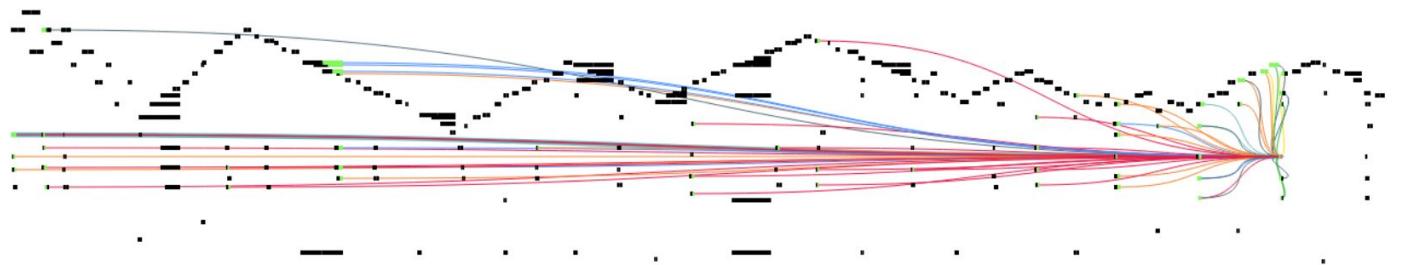
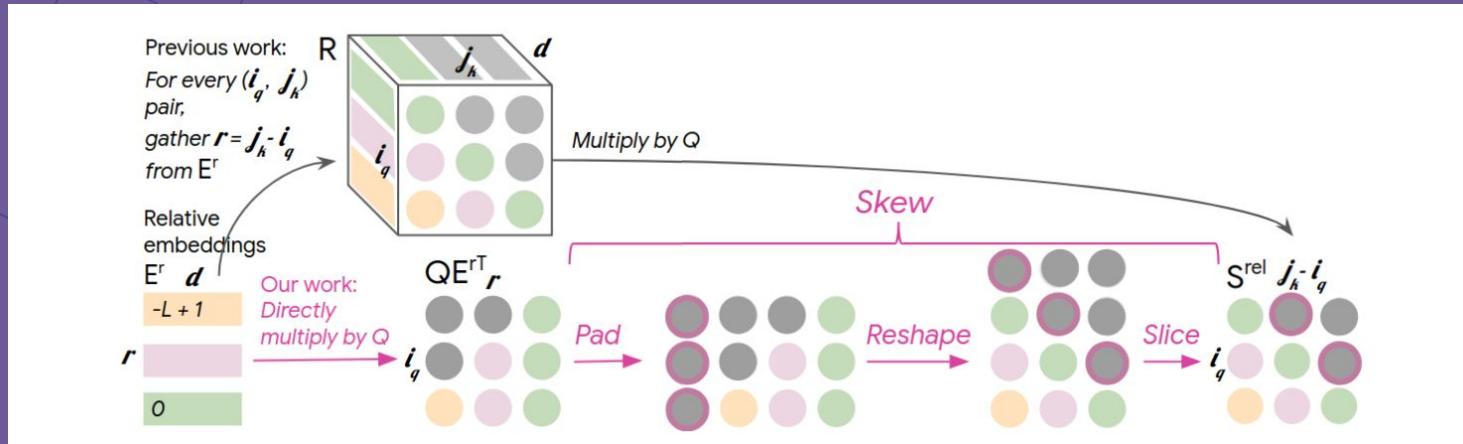


Figure 9: The query a note in the left-hand, and it attends to its immediate past neighbors and mostly to the earlier left hand chords, with most attention lines distributed in the lower half of the pianoroll.

# Memory Efficient implementation Of Relative Position-Based Attention



# Memory Efficient implementation Of Relative Position-Based Attention

Table 1: Comparing the overall relative memory complexity (intermediate relative embeddings ( $R$  or  $E^r$ ) + relative logits  $S^{rel}$ ), the maximal training lengths that can fit in a GPU with 16GB memory assuming  $D_h = 64$ , and the memory usage per layer per head (in MB).

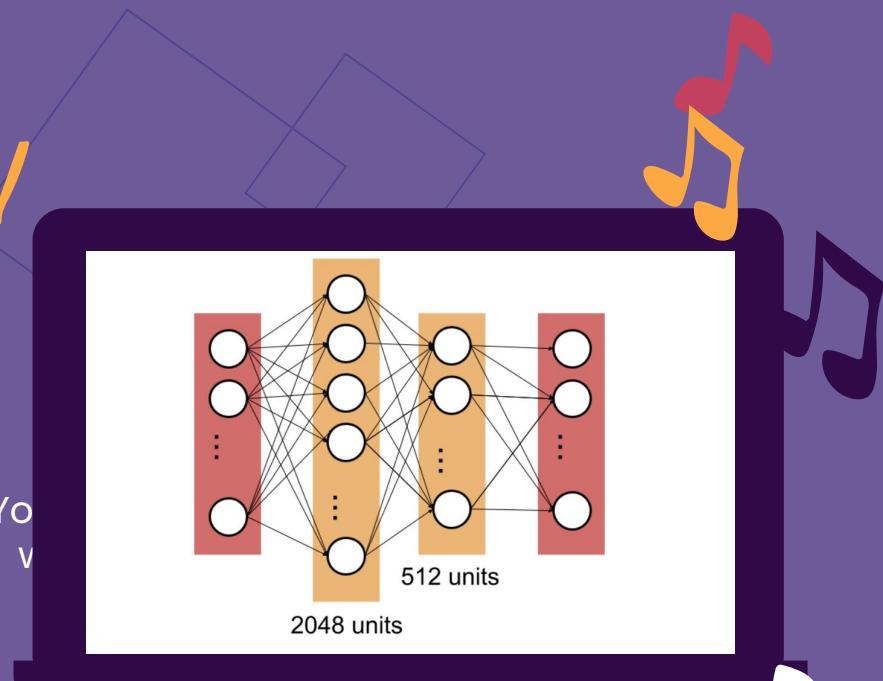
Implementation	Relative memory	Maximal $L$	$L = 650$	$L = 2048$	$L = 3500$
Shaw et al. (2018)	$O(L^2D + L^2)$	650	108 + 1.7	1100 + 16	3100 + 49
Ours	$O(LD + L^2)$	3500	0.17 + 1.7	0.52 + 16	0.90 + 49

# Feedforward

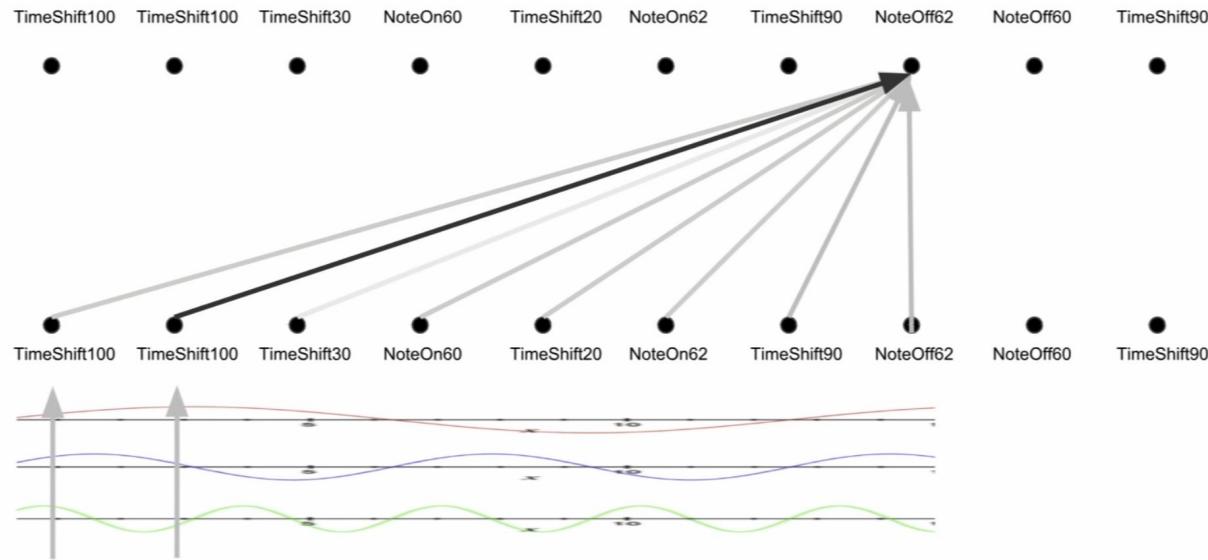
$$\text{FF}(Z) = \text{ReLU}(ZW_1 + b_1)W_2 + b_2$$

$$f(x) = \max(0, x) = \begin{cases} 0 & \text{for } x \leq 0 \\ x & \text{for } x > 0 \end{cases}$$

2048 units  
512 units

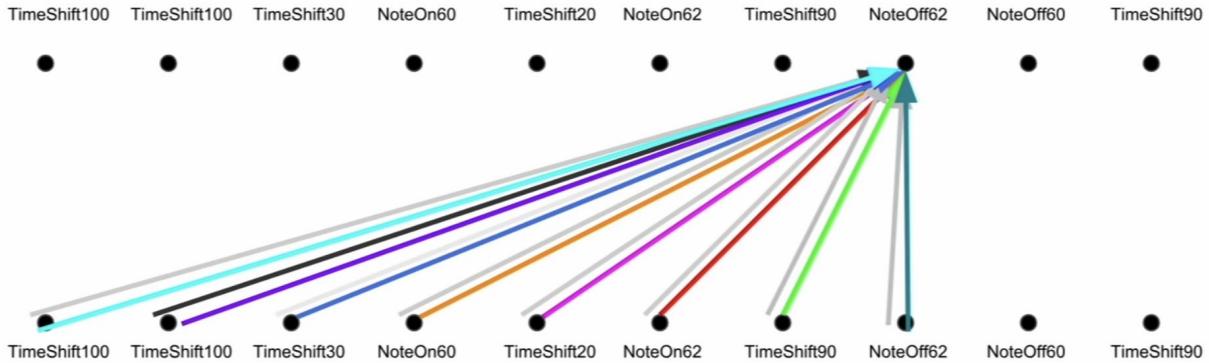


# Attention: a weighted average



# Relative attention (Shaw et al, 2018)

## Multihead attention + convolution?



## D PREVIOUS FIGURES FOR THE “SKEWING” PROCEDURE

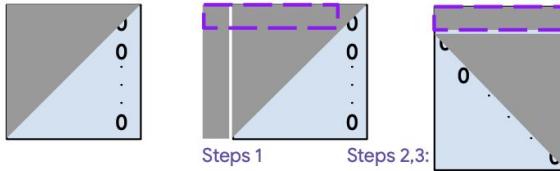


Figure 10: Relative global attention: Steps (from left to right) for “skewing” an absolute-by-relative  $(i_q, r)$  indexed matrix into absolute-by-absolute  $(i_q, j_k)$ . Grey indicates self-attention masks or entries introduced by the skewing procedure. Positions with relative distance zero are marked. Entries outlined by purple are removed in step 3.

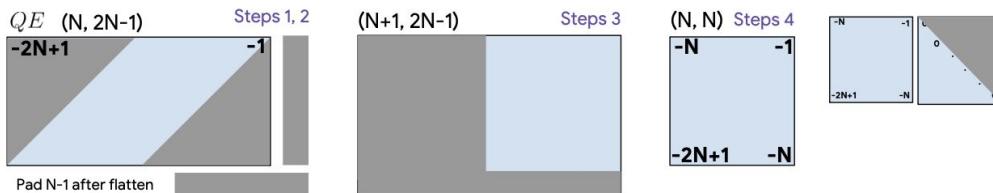


Figure 11: Relative local attention: Steps (from left to right) for “skewing” an  $(i_q, r)$  indexed matrix with  $2N - 1$  ranged relative indices  $r$  into  $(i_q, j_k)$  indexed. Shapes are indicated above the boxes, while indices in the boxes give relative distances.

# Word Vector

Key  
 $K = XW^K$

Query  
 $Q = XW^Q$



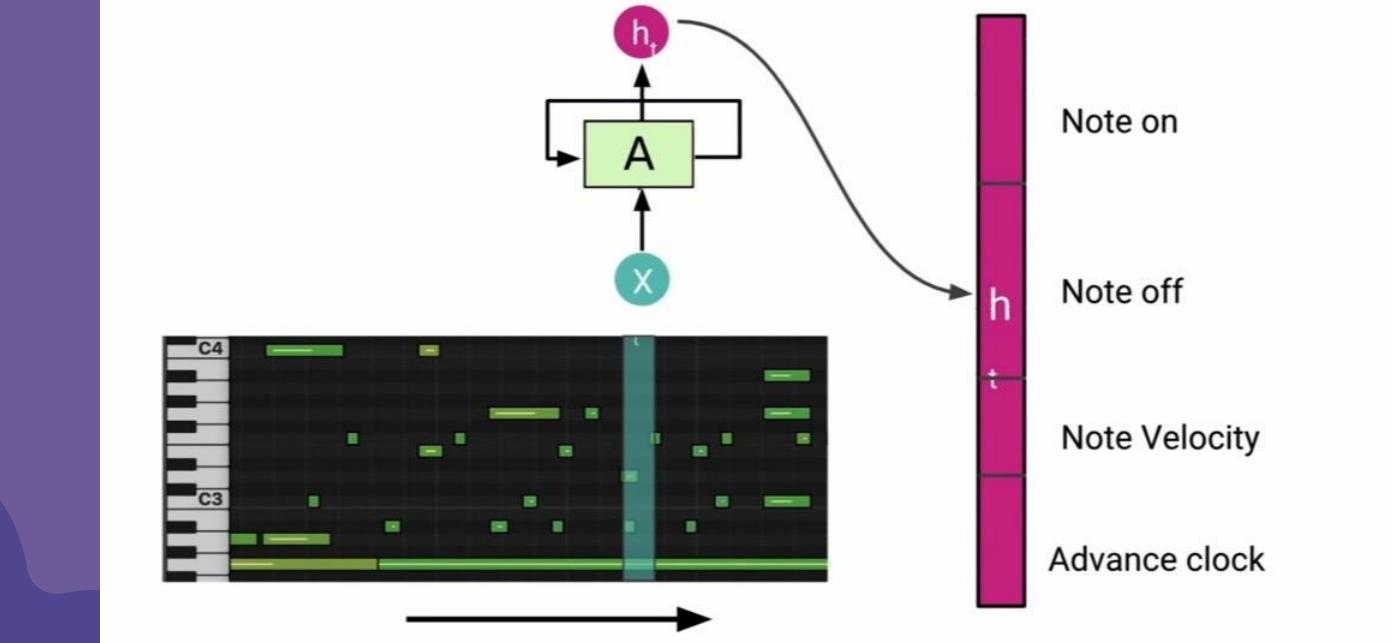
Value

$$V = XW^V$$

Vector

$$X = (x_1, x_2, x_3, \dots, x_L)$$

# Music Language model: Prior work Performance RNN



$$FF(Z) = \text{ReLU}(ZW_1 + b_1)W_2 + b_2$$

$$S^{\text{rel}} = QR^T$$

**CREDITS:** This presentation template was created by [Slidesgo](#), including icons by [Flaticon](#) and infographics & images by [Freepik](#)

# Thanks!

Do you have any questions?

[youremail@freepik.com](mailto:youremail@freepik.com)  
+91 620 421 838  
[yourcompany.com](http://yourcompany.com)

**CREDITS:** This presentation template was created by [Slidesgo](#), including icons by [Flaticon](#) and infographics & images by [Freepik](#)



# World Saxophone Day

Here is where your presentation begins!



# Contents of this template

Here's what you'll find in this **Slidesgo** template:

1. A slide structure based on a Multi-purpose presentation, which you can easily adapt to your needs. For more info on how to edit the template, please visit **Slidesgo School** or read our **FAQs**.
2. An assortment of graphic resources that are suitable for use in the presentation can be found in the **alternative resources slide**.
3. A **thanks slide**, which you must keep so that proper credits for our design are given.
4. A **resources slide**, where you'll find links to all the elements used in the template.
5. **Instructions for use**.
6. Final slides with:
  - The **fonts and colors** used in the template.
  - A selection of **illustrations**. You can also customize and animate them as you wish with the online editor. Visit **Storyset** to find more.
  - More **infographic resources**, whose size and color can be edited.
  - Sets of **customizable icons** of the following themes: general, business, avatar, creative process, education, help & support, medical, nature, performing arts, SEO & marketing, and teamwork.

You can delete this slide when you're done editing the presentation

# Contents of this template

Here's what you'll find in this **Slidesgo** template:

1. A slide structure based on a Multi-purpose presentation, which you can easily adapt to your needs. For more info on how to edit the template, please visit **Slidesgo School** or read our **FAQs**.
2. An assortment of graphic resources that are suitable for use in the presentation can be found in the **alternative resources slide**.
3. A **thanks slide**, which you must keep so that proper credits for our design are given.
4. A **resources slide**, where you'll find links to all the elements used in the template.
5. **Instructions for use**.
6. Final slides with:
  - The **fonts and colors** used in the template.
  - A selection of **illustrations**. You can also customize and animate them as you wish with the online editor. Visit **Storyset** to find more.
  - More **infographic resources**, whose size and color can be edited.
  - Sets of **customizable icons** of the following themes: general, business, avatar, creative process, education, help & support, medical, nature, performing arts, SEO & marketing, and teamwork.

You can delete this slide when you're done editing the presentation

# Table of contents

01

## The instrument

You can describe the topic  
of the section here

03

## Future plans

You can describe the topic  
of the section here

02

## Some history

You can describe the topic  
of the section here

04

## Social media

You can describe the topic  
of the section here



# Whoa!

Saxophone Day is a celebration of Adolphe Sax's birth date and contributions to the music business!



# Jntroduction

Every year on November 6th, this day is commemorated to mark the birth anniversary of Adolphe Sax, who was born on that date in 1814. Mr. Sax is the creator of the saxophone, one of the most important instruments in jazz music



01

# The instrument

You can enter a subtitle here if you need it

# Some interesting facts

Do you know what helps you make your point clear?  
Lists like this one:

- They're simple
- You can organize your ideas clearly
- You'll never forget to buy milk!

And the most important thing: the audience won't miss the point of your presentation



# Where it's used



## Theaters

Mercury is the closest planet to the Sun and the smallest one in the Solar System



## Festivals

Venus has a beautiful name and is the second planet from the Sun



## Concerts

Despite being red, Mars is actually a cold place. It's full of iron oxide dust

# Next events

## June

Despite being red,  
Mars is actually a  
very cold place



## July

Neptune is the  
farthest planet  
from the Sun

## August

Venus is the  
second planet  
from the Sun



## September

Saturn is a gas  
giant and has  
several rings



# Best companions



## Guitar

Venus is the second planet from the Sun



## Piano

Jupiter is the biggest planet in our System



## Violin

Despite being red, Mars is a cold place



## Banjo

Saturn is a gas giant and has several rings



# Different types

## Tenor

Mercury is the closest planet to the Sun

## Soprano

Jupiter is the biggest planet of them all

## Baritone

Venus is the second planet from the Sun

## Bass

Saturn is composed of hydrogen and helium

## Alto

Mars is actually a very cold place

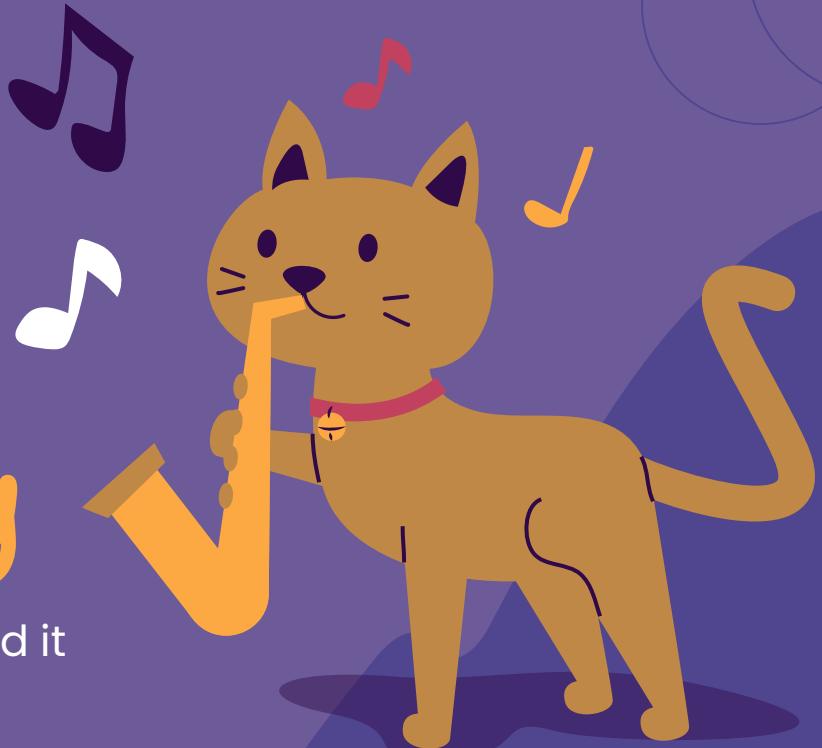
## Tubax

Neptune is the farthest planet from the Sun

02

# Some history

You can enter a subtitle here if you need it



J'm never  
too loud





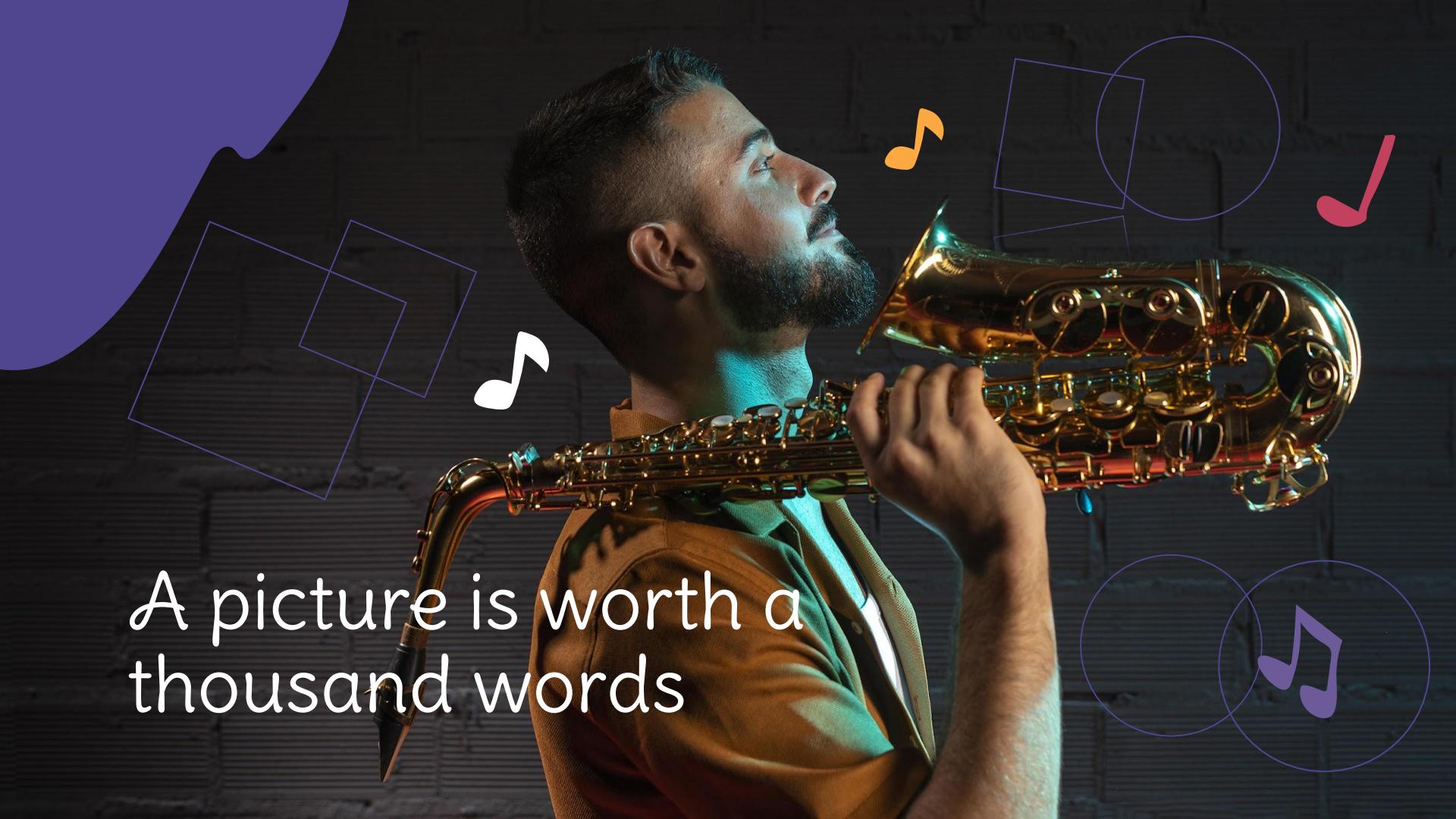
“Don’t play the saxophone.  
Let it play you”

—Charlie Parker

# This is a great headline

And this is the subtitle that makes it comprehensible





A picture is worth a  
thousand words

# A picture always reinforces the concept

Images reveal large amounts of data, so remember: use an image instead of a long text. Your audience will appreciate it





03

# Future plans

You can enter a subtitle here if you need it

# Important percentages

40%



Mars

Despite being red, Mars is actually a cold place. It's full of iron oxide dust

20%



Venus

Venus has a beautiful name and is the second planet from the Sun

60%



Mercury

It's the closest planet to the Sun and the smallest one in our System

# Let's compare sounds

## Analog



Venus has a beautiful name and is the second planet from the Sun

## Digital



Despite being red, Mars is a cold place. It's full of iron oxide dust



# Did you know?

Although the saxophone is most commonly associated with jazz, it has also been employed successfully in symphonic works like Berlioz, Bizet, and Massenet!

04

# Social media

You can enter a subtitle here if you need it



# Laptop software

You can replace the image on the screen with your own work. Just right-click on it and select "Replace image"



# Tablet app

You can replace the image on the screen with your own work. Just right-click on it and select “Replace image”



# Smartphone

You can replace the image on the screen with your own work. Just right-click on it and select “Replace image”



# Our team



## James Doe

You can speak a bit  
about this person here



## Timmy Jimmy

You can speak a bit  
about this person here



## Susan Bones

You can speak a bit  
about this person here

# Where it's used the most



60%

- **Classic jazz**

Venus is the second planet from the Sun

40%

- **Blues**

Saturn is a gas giant with several rings



# Our courses schedule



Week 1

Week 2

Week 3

Practice 1

8:30 a.m.

9:30 a.m.

10:30 a.m.

Practice 2

9:30 a.m.

10:30 a.m.

11:30 a.m.

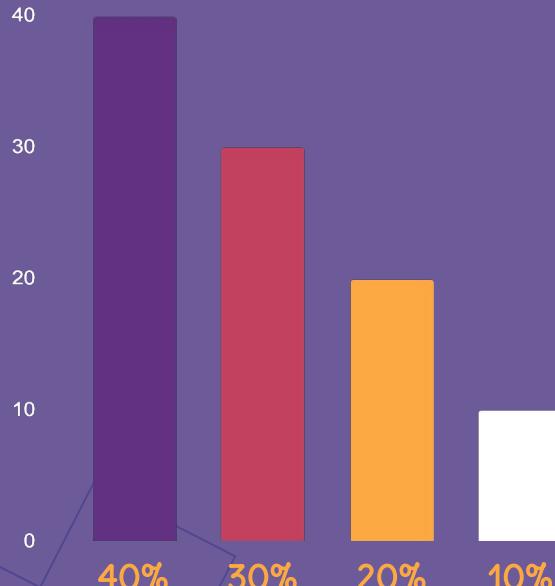
Practice 3

12:30 p.m.

1:30 p.m.

2:30 p.m.

# Check this graph



Follow the link in the graph to modify its data and then paste the new one here. **[For more info, click here](#)**

## Venus

Venus is the second planet from the Sun

## Mars

Despite being red, Mars is actually cold

## Saturn

Saturn is a gas giant and has several rings

## Neptune

Neptune is the farthest planet from the Sun

# Conclusions

## Mercury

Mercury is the closest planet to the Sun



## Jupiter

Jupiter is the biggest planet of them all



## Venus

Venus is the second planet from the Sun



## Saturn

It's composed of hydrogen and helium



# Thanks!

Do you have any questions?

[youremail@freepik.com](mailto:youremail@freepik.com)  
+91 620 421 838  
[yourcompany.com](http://yourcompany.com)



**CREDITS:** This presentation template was created by [Slidesgo](#), including icons by [Flaticon](#) and infographics & images by [Freepik](#)

Please keep this slide for attribution



# Jcon Pack: Jazz Music



# Alternative resources

Here's an assortment of alternative resources whose style fits the one of this template

- Hand drawn international jazz day vertical poster template
- Different pets illustration concept
- Different pets set



# Resources

Did you like the resources on this template? Get them for free at our other websites:



## Vectors

- Hand drawn international jazz day illustration
- Organic flat international jazz day vertical poster template
- Hand drawn international jazz day illustration
- Different pets pack
- Different pets collection
- Hand drawn international jazz day illustration
- Flat international jazz day illustration

## Photos

- Front view of man playing the saxophone
- Side view of handsome male musician holding saxophone
- Close up man playing the guitar
- Portrait female with guitar
- Side view of male musicians at home playing guitar

## Icons

- Icon Pack: Jazz Music

# Instructions for use

In order to use this template, you must credit **Slidesgo** by keeping the **Thanks** slide.

**You are allowed to:**

- Modify this template.
- Use it for both personal and commercial projects.

**You are not allowed to:**

- Sublicense, sell or rent any of Slidesgo Content (or a modified version of Slidesgo Content).
- Distribute Slidesgo Content unless it has been expressly authorized by Slidesgo.
- Include Slidesgo Content in an online or offline database or file.
- Offer Slidesgo templates (or modified versions of Slidesgo templates) for download.
- Acquire the copyright of Slidesgo Content.

For more information about editing slides, please read our FAQs or visit Slidesgo School:

<https://slidesgo.com/faqs> and <https://slidesgo.com/slidesgo-school>

# Instructions for use (premium users)

As a Premium user, you can use this template without attributing Slidesgo or keeping the "Thanks" slide.

## You are allowed to:

- Modify this template.
- Use it for both personal and commercial purposes.
- Hide or delete the "Thanks" slide and the mention to Slidesgo in the credits.
- Share this template in an editable format with people who are not part of your team.

## You are not allowed to:

- Sublicense, sell or rent this Slidesgo Template (or a modified version of this Slidesgo Template).
- Distribute this Slidesgo Template (or a modified version of this Slidesgo Template) or include it in a database or in any other product or service that offers downloadable images, icons or presentations that may be subject to distribution or resale.
- Use any of the elements that are part of this Slidesgo Template in an isolated and separated way from this Template.
- Register any of the elements that are part of this template as a trademark or logo, or register it as a work in an intellectual property registry or similar.

For more information about editing slides, please read our FAQs or visit Slidesgo School:

<https://slidesgo.com/faqs> and <https://slidesgo.com/slidesgo-school>

# Fonts & colors used

This presentation has been made using the following fonts:

## **Delius Swash Caps**

(<https://fonts.google.com/specimen/Delius+Swash+Caps>)

## **Poppins**

(<https://fonts.google.com/specimen/Poppins>)

#300948

#623182

#50468f

#6d5b99

#a6304c

#c2415f

#bf8847

#fca943

#e5e5e5

#ffffff

# Storyset

Create your Story with our illustrated concepts. Choose the style you like the most, edit its colors, pick the background and layers you want to show and bring them to life with the animator panel! It will boost your presentation. Check out [How it Works](#).



Pana



Amico



Bro



Rafiki



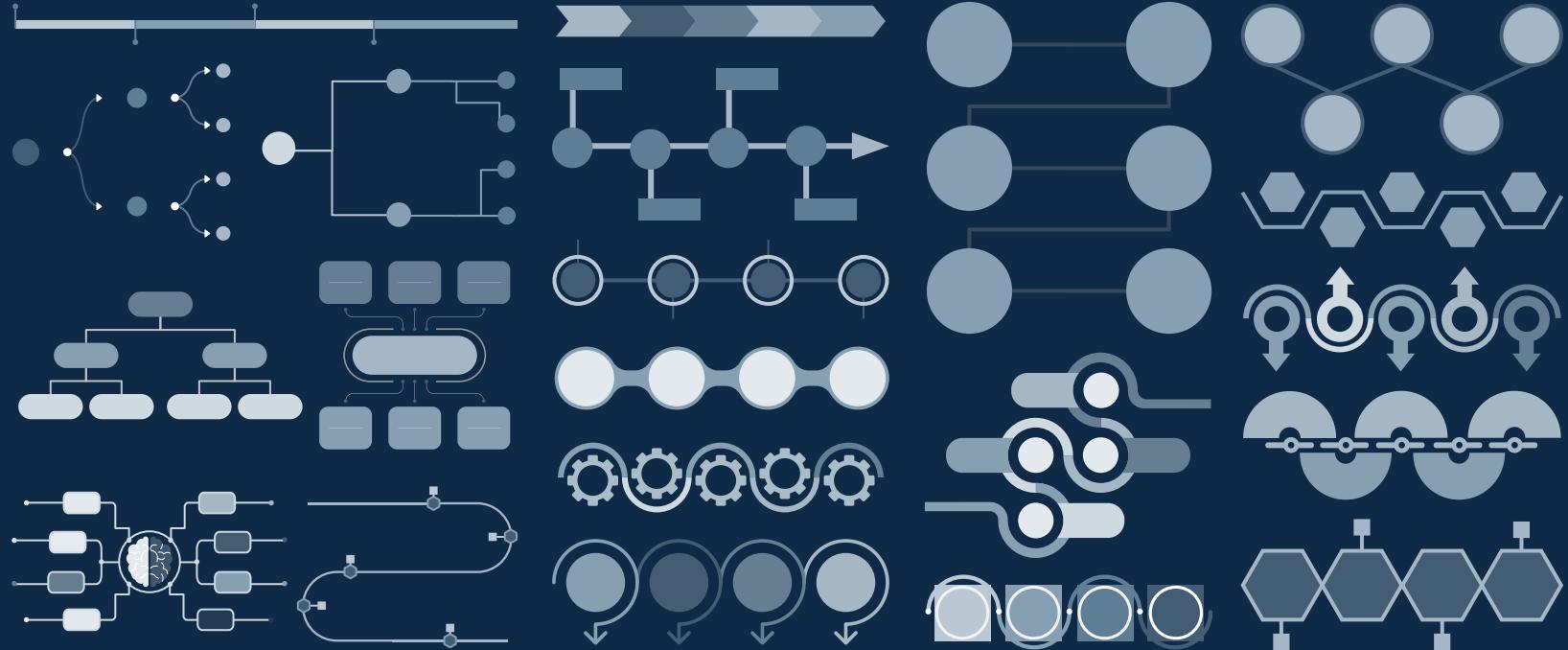
Cuate

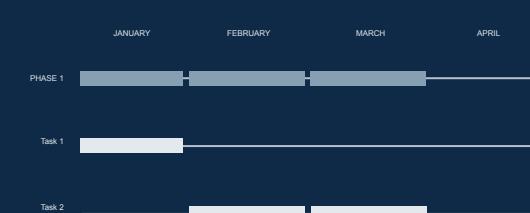
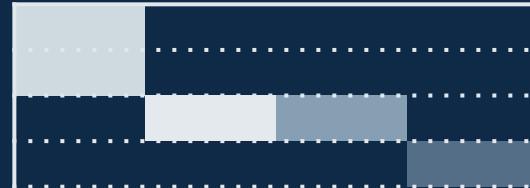
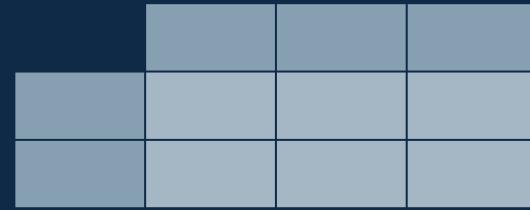
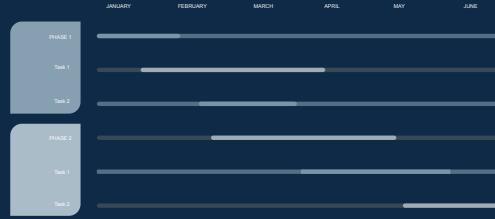
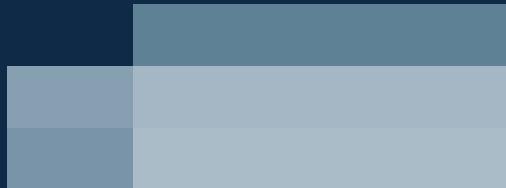
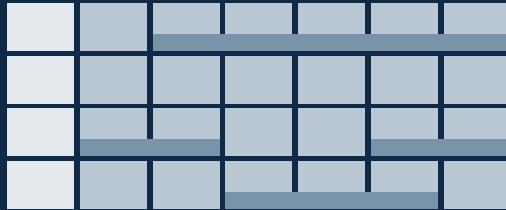
# Use our editable graphic resources...

You can easily resize these resources without losing quality. To change the color, just ungroup the resource and click on the object you want to change. Then, click on the paint bucket and select the color you want. Group the resource again when you're done. You can also look for more infographics on Slidesgo.

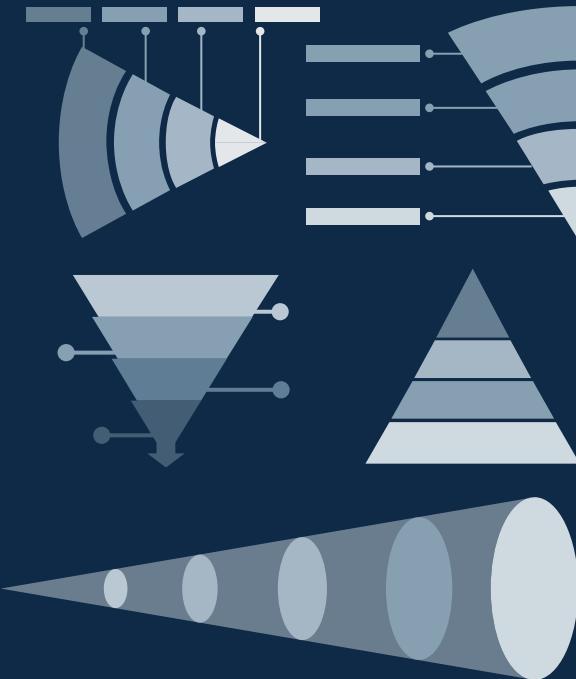
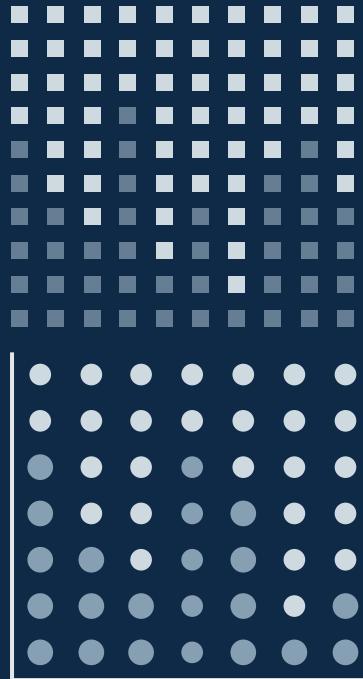












# ...and our sets of editable icons

You can resize these icons without losing quality.

You can change the stroke and fill color; just select the icon and click on the paint bucket/pen.

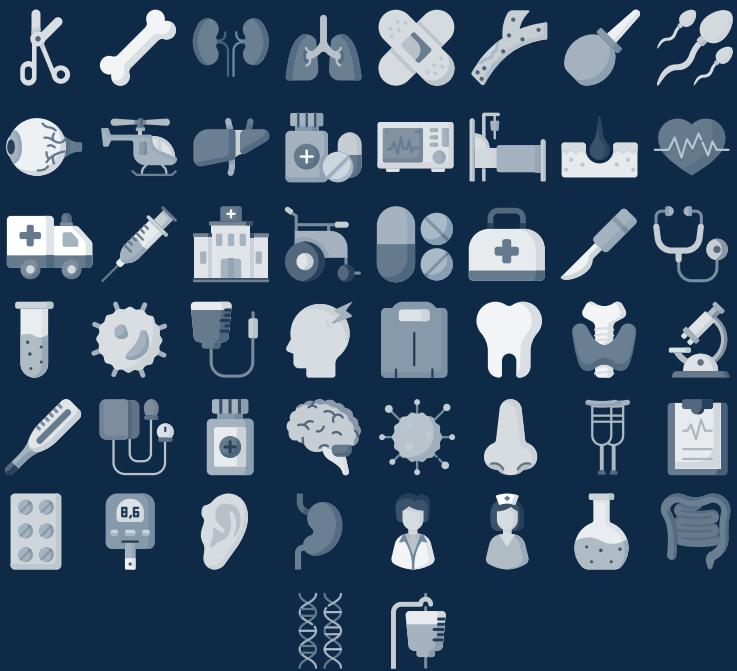
In Google Slides, you can also use Flaticon's extension, allowing you to customize and add even more icons.



## Educational Icons



## Medical Icons



## Business Icons



## Teamwork Icons



## Help & Support Icons



# Avatar Icons



## Creative Process Icons



## Performing Arts Icons



# Nature Icons



# SEO & Marketing Icons



