

# Neural Monkey: The Current State and Beyond

Jindřich Helcl, Jindřich Libovický, Tom Kocmi,  
Tomáš Musil, Ondřej Cífka, Dušan Variš, Ondřej Bojar

Institute of Formal and Applied Linguistics  
Faculty of Mathematics and Physics  
Charles University

March 19, 2018



# Neural Monkey

---

- Open-source toolkit for sequential learning
- Suited for research and education
- Three (overlapping) user groups considered:
  - Students
  - Researchers
  - Newcomers to deep learning

# Goals

---

1. Code readability

# Goals

---

1. Code readability
2. Modularity along research concepts

## Goals

---

1. Code readability
2. Modularity along research concepts
3. Up-to-date building blocks

# Goals

---

1. Code readability
2. Modularity along research concepts
3. Up-to-date building blocks
4. Fast prototyping

## Development

---

- Implemented in Python 3.5 using TensorFlow
- thanks to TensorFlow GPU support using CUDA, cuDNN
- Actively developed using GitHub as the main communication platform



Source code here:

<https://github.com/ufal/neuralmonkey>





# Simple MT Configuration Example

---

## [main]

```
output="output_dir"
batch_size=64
epochs=20
train_dataset=<train_data>
val_dataset=<val_data>
trainer=<my_trainer>
runners=[<my_runner>]
evaluation=[
    ("target", evaluators.BLEU)]
logging_period=500
validation_period=5000
```

## [en\_vocabulary]

```
class=vocabulary.from_wordlist
path="en_vocab.tsv"
```

## [de\_vocabulary]

```
class=vocabulary.from_wordlist
path="de_vocab.tsv"
```

## [train\_data]

```
class=dataset.from_files
s_source="data/train.en"
s_target="data/train.de"
```

## [val\_data]

```
class=dataset.from_files
s_source="data/val.en"
s_target="data/val.de"
```

## [my\_encoder]

```
class=encoders.SentenceEncoder
rnn_size=500
embedding_size=600
data_id="source"
vocabulary=<en_vocabulary>
```

## [my\_attention]

```
class=attention.Attention
encoder=<my_encoder>
state_size=500
```

## [my\_decoder]

```
class=decoders.Decoder
encoders=[<my_encoder>]
attentions=[<my_attention>]
rnn_size=1000
embedding_size=600
data_id="target"
vocabulary=<de_vocabulary>
```

## [my\_trainer]

```
class=trainers.CrossEntropyTrainer
decoders=[<my_decoder>]
clip_norm=1.0
```

## [my\_runner]

```
class=runners.GreedyRunner
decoder=<my_decoder>
```

# Configuration of Captioning

---

Define how to load images:

```
[imagenet_reader]
class=readers.image_reader.imagenet_reader
prefix="/lnet/troja/projects/wmt17-multimodal/data/flickr30k-images"
target_width=229
target_height=229
zero_one_normalization=True
```

And replace encoder with ImageNet network:

```
[imagenet_resnet]
class=encoders.ImageNet
name="imagenet"
data_id="images"
network_type="resnet_v2_50"
spatial_layer="resnet_v2_50/block4/unit_3/bottleneck_v2/conv3"
slim_models_path="tensorflow-models/research/slim"
```

# Sentence Classification

---

Keep the encoder and replace the decoder (and update the rest)

[my\_classifier]

```
class=decoders.Classifier
data_id="labels"
encoders=[<my_encoder>]
vocabulary=<label_vocabulary>
layers=[200]
```

[my\_runner]

```
class=runners.PlainRunner
decoder=<my_classifier>
```

[my\_trainer]

```
class=trainers.CrossEntropyTrainer
decoders=[<my_classifier>]
clip_norm=1.0
```

## Pre-trained model parts

Parameters of model parts can be fixed using the gradient blocking module

# Supported Features

---

- Recurrent encoder and decoder with attention
- Beam search decoding with model ensembling
- Deep convolutional encoder
- Self-attentive encoder and decoder
- Wrappers for ImageNet networks
- Custom CNNs for image processing
- ConvNets for sequence classification
- Self-attentive embeddings for sentence classification
- Hierarchical attention over multiple source sequences
- Generic sequence labeler
- Connectionist temporal classification

# Console Logging during Training

```
[11]
source: it started the French League season with two wins and two draws , and on Sunday it defeated St. Etienne 5 : 0 .
source_bpe: it started the French League season with two wins and two draws , and on Sunday it defeated St. Eti@enne 5 : 0 .
target_bpe: ten do se@@ z@@ ony francouzské ligy od@@ stat@@ oval dvěma výhra@@ mi a dvěma re@@ mi@@ zami , v neděli roz@@ střílel St . Eti@enne 5 : 0 .
target: začalo to <unk> <unk> <unk> <unk> a <unk> , a v neděli <unk> <unk> <unk> .
target (ref): ten do sezony francouzské ligy odstatoval dvěma výhrami a dvěma remizami , v neděli rozstřílel St . Etienne 5 : 0 .

[12]
source: it is putting the most pressure on AS Monaco .
source_bpe: it is putting the most pressure on AS Monaco .
target_bpe: relativně nejvíc tlačí bota klub AS Monaco .
target: <unk> <unk> tlak na <unk> AS .
target (ref): relativně nejvíc tlačí bota klub AS Monaco .

[13]
source: it so far has a balance of 1-1-2 in the league , and over the weekend it parted with Lille with a 1 : 1 tie .
source_bpe: it so far has a balance of 1-@@ 1-2 in the league , and over the weekend it parted with L@@ ille with a 1 : 1 tie .
target_bpe: ten má v lize zatím bilanci 1-@@ 1-2 a o víkendu se rozešel s L@@ ille s@@ mírně 1 : 1 .
target: v <unk> <unk> <unk> <unk> a <unk> se <unk> s <unk> <unk> <unk> .
target (ref): ten má v lize zatím bilanci 1-1-2 a o víkendu se rozešel s Lille smírně 1 : 1 .

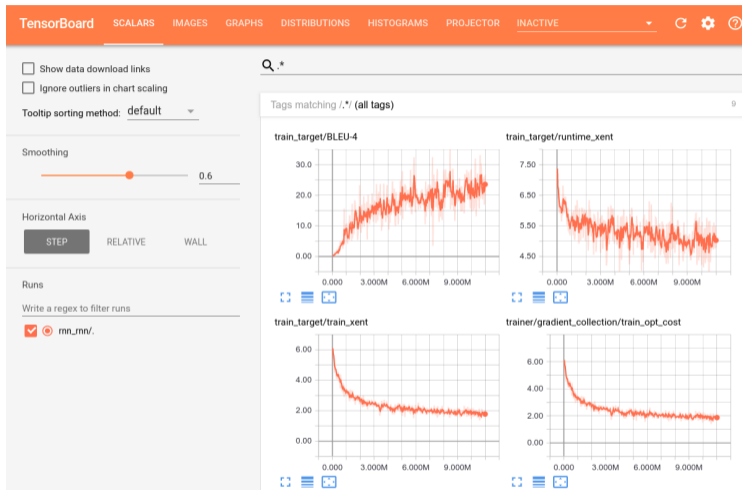
[14]
source: today news appeared on tabloid English-language websites stating that the Queens Park Rangers may want Čech .
source_bpe: today news appeared on ta@@ blo@@ id Engli@@ sh-@@ language websites stating that the Queens Park Rangers may want Č@@ ech .
target_bpe: na bulvá@@ rních anglic@@ kých we@@ bch dnes vy@@ sko@@ čily zprávy , že by Č@@ ech@@ a mohli chtít v Queens Park Rangers .
target: dnes se objevila zpráva na internetových stránkách <unk> <unk> , které <unk> , že <unk> <unk> <unk> .
target (ref): na bulvárních anglických webech dnes vyskočily zprávy , že by Čecha mohli chtít v Queens Park Rangers .

[15]
source: however , according to English bookmakers , Čech &apos;s departure from Chelsea at the last minute is not very likely .
source_bpe: however , according to English book@@ makers , Č@@ ech &@@ a@@ pos@@ ;@@ s departure from Chelsea at the last minute is not very likely .
target_bpe: podle anglic@@ kých book@@ ma@@ ker@@ ů ale Č@@ ech@@ ův odchod z Chelsea na poslední chvíli není příliš pravděpodobný .
target: nicméně , podle <unk> <unk> , <unk> <unk> <unk> z Chelsea za poslední minutu není pravděpodobné .
target (ref): podle anglických bookmakerů ale Čechův odchod z Chelsea na poslední chvíli není příliš pravděpodobný .

2018-03-14 12:58:13: Validation (epoch 1, batch number 214960):
2018-03-14 12:58:46: Variable file saved in experiments/czeng_rnn/variables.data
2018-03-14 12:58:46: Best scores saved so far: [9.8549028005205079]
2018-03-14 12:59:06: Variables of 'encoder' saved to 'experiments/czeng_rnn/encoder.ckpt'
2018-03-14 12:59:26: Variables of 'encoder_input' saved to 'experiments/czeng_rnn/embeddings.ckpt'
2018-03-14 12:59:26: best target/BLEU-4 on validation: 9.855 (in epoch 1, after batch number 214960)
2018-03-14 12:59:26: Epoch 1/10 Instances 10748050 target/runtime_xent: 6.395 target/train_xent: 2.626 target/BLEU-4: 9.855
2018-03-14 12:59:26: Validation time: 11713.46s, inter-validation: 7391.83s, per-instance (train): 0.02s, per-instance (val): 4.41s
2018-03-14 12:59:26: Validation period setting is inefficient.

2018-03-14 13:03:11: Epoch 1/10 Instances 10748100 target/runtime_xent: 5.032 target/train_xent: 1.726 target/BLEU-4: 20.54
2018-03-14 13:15:41: Epoch 1/10 Instances 10778450 target/runtime_xent: 5.367 target/train_xent: 2.043 target/BLEU-4: 20.02
2018-03-14 13:28:07: Epoch 1/10 Instances 10808500 target/runtime_xent: 4.734 target/train_xent: 1.885 target/BLEU-4: 18.73
2018-03-14 13:40:35: Epoch 1/10 Instances 10838550 target/runtime_xent: 5.606 target/train_xent: 1.766 target/BLEU-4: 23.88
2018-03-14 13:53:03: Epoch 1/10 Instances 10868350 target/runtime_xent: 5.246 target/train_xent: 1.713 target/BLEU-4: 24.79
```

# Scalar Values in TensorBoard



Losses, evaluation metrics, parameter norms, histograms of gradients

# Attention in TensorBoard

TensorBoard SCALARS IMAGES GRAPHS DISTRIBUTIONS HISTOGRAMS PROJECTOR INACTIVE

Show actual image size

Runs

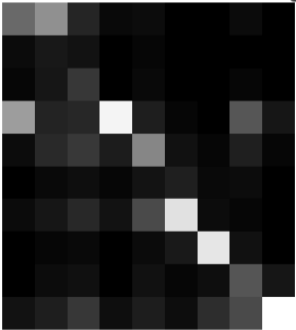
Write a regex to filter runs

- san\_ctc/
- rnn\_rnn/
- san\_ctc\_proj/

Q .\*

Tags matching /.\*/ (all tags) 1

decoder/attention/image/0 rnn\_rnn/  
step 10,748,050 Wed Mar 14 2018 12:59:26 GMT+0100 (CET)



# Conclusions

---

Neural Monkey is:

- Actively developing open-source GitHub project
- Suited for researchers, students, and other DL enthusiasts
- Large collection of interesting features from across the NLP sub-topics
- Simple tool to use because of clear and readable configuration files
- Highly modular, therefore also relatively easy to debug



# Conclusions

---

Neural Monkey is:

- Actively developing open-source GitHub project
- Suited for researchers, students, and other DL enthusiasts
- Large collection of interesting features from across the NLP sub-topics
- Simple tool to use because of clear and readable configuration files
- Highly modular, therefore also relatively easy to debug

Thank you for your attention!