Summarizing Source Code using a Neural Attention Model

Srinivasan Iyer, Ioannis Konstas, Alvin Cheung, Luke Zettlemoyer
University of Washington, Seattle, USA

Motivation

We present the first completely data driven method to generate high level summaries of the function of code.

C#
```
var input = "Hello";
var regEx = new Regex("World");
return !regEx.IsMatch(input);
```
→ Lookup a substring in a string using regex

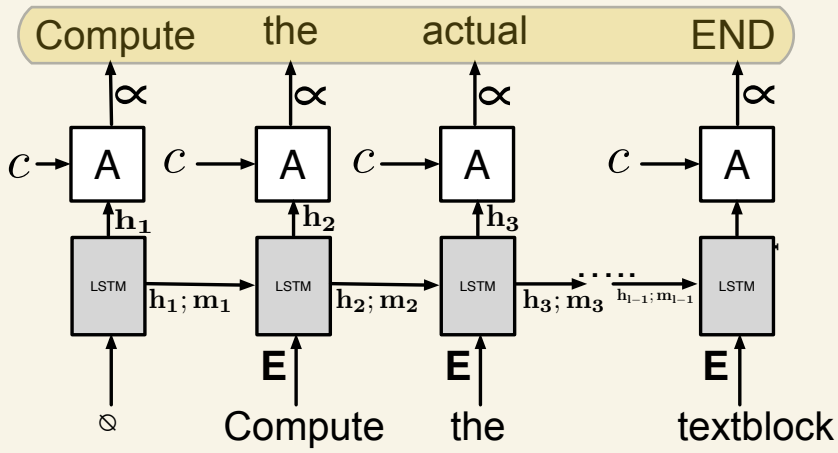
SQL
```
SELECT MAX(marks) FROM ATable
WHERE marks <
 (SELECT MAX(marks) FROM ATable)
```
→ Get the second largest value of a column

These auto-generated summaries have many Software Engineering applications:

Code Search · Documentation · Code Folding · Code Retrieval

Neural Attention Model

We use an end-to-end model that jointly performs content selection using an attention mechanism, and surface realization using Long Short Term Memory networks.

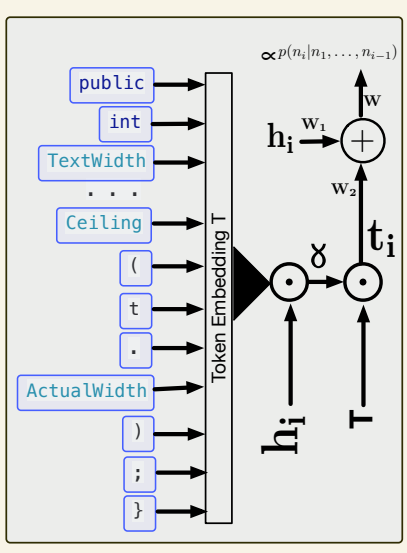


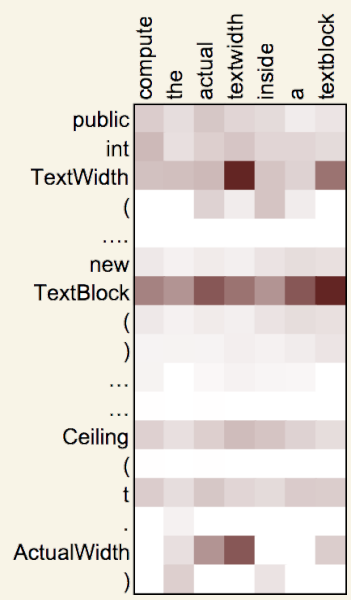
We model the conditional next-word probability as:

$$p(n_i|n_1,\ldots,n_{i-1}) \propto \mathbf{W}\ \tanh(\mathbf{W_1 h_i} + \mathbf{W_2 t_i})$$

$\mathbf{h_i}$ is the hidden state of the LSTM cell at the time step **i**

The attention model computes a weighted sum $\mathbf{t_i}$ of the token embeddings based on the LSTM hidden state. In this way, it selects the most useful tokens to generate the current word.





Code Summarization Dataset

NEW

We create a new dataset from programming QA websites containing 66K examples for C# and 33K for SQL.

stackoverflow

Compute the actual textwidth inside a textblock

1 Answer — active **oldest** votes

5

```
public int TextWidth(string text) {
  TextBlock t = new TextBlock();
  t.Text = text;
  return (int)Math.Ceiling(t.ActualWidth);
}
```

C#

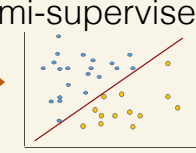
MSDN about ActualWidth.

share edit — answered Dec 31 '13 at 13:44

Code snippets in this dataset are non-trivial.

Loops	> 20%	> 2 Functions	50%	|Code|	38
Conditionals	> 22%	> 2 Statements	45%	|Summary|	12

Titles are cleaned using an semi-supervised classifier.

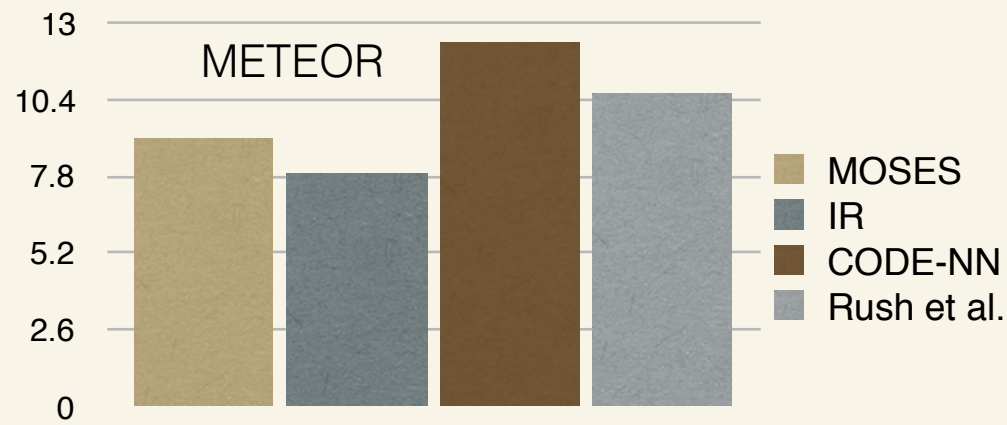
Difficult C# if then logic →  REJECTED

Human Annotations

We gather 2 additional references for 200 code snippets for more accurate development and testing.
Dataset/Code at: https://github.com/sriniiyer/codenn

Experiments

Our model beats competitive baselines on summarization metrics such as METEOR and BLEU-4.



Human evaluators prefer the output of our model too!