Moses Version 4.0 Release Notes

Overview

The document contains the release notes for the Moses SMT toolkit, version 4.0. It describes the changes to the toolkit since version 3.0 from January 2015 to September 2017.

Machine translation research has shifted to using neural networks. While there is still research on advancing the models implemented in Moses, there has been a noticeable reduction in the amount of communication on the mailing list and contribution to Moses codebase. The toolkit is still widely used as the de-facto SMT baseline and piecemeal as pre-processing and post-processing for neural MT and other research. And of course, it is still widely used in production by organizations that have invested a lot of time and effort into Moses.

This release will be the last fully tested release for the forseeable future. The aim of this release is ensure that Moses continues to function and will do so in the future where there is a reduced level of support from the MT community. Therefore, we have endeavoured to test the end-to-end training and translation pipeline on a number of up-to-date distributions and all major operating systems.

Of course, we have also tested major new additions to the Moses toolkit such as Implz (a fast, memory efficient language model trainer), the OSM feature function for best phrase-based translation, and Moses2 (a fast multi-threaded decoder).

We make the available for download the executable for Linux (Ubuntu 16.04 and 17.04), macOS (Sierra) and Windows 10 with Ubuntu 16.04 subsystem. All also make available the trained model and training configuration files in order to replicate our experiments.

New Features

The following is a list of the major new features in the Moses toolkit since version 3.0, in roughly reverse chronological order.

1. RDLM, a language model for the string-to-tree decoder with a dependency grammar, by Rico Sennrich.
2. Docker files for Moses, by Liling Tan
3. Config files for Travis CI (code testing framework), by sih4sing5hong5
4. Moses2 (fast Moses-compatible decoder for multi-core servers), by Hieu Hoang
5. Moses2 compilation with Visual Studio on Windows
6. Chinese non-breaking prefix files, by Linas Vepštas

7. Per sentence language model and phrase table, by Lane Schwartz
8. ProbingPT, a fast phrase-table implementation for Moses and Moses2, by Nikolay Bogoychev
9. Operation sequence model, for better quality phrase-based decoding, by Nadir Durrani
10. Compile with C++11 by default
11. SoftSourceSyntacticConstraintsFeature, RulePairUnlexicalizedSource, Model1Feature, PhraseOrientationFeature, SoftSourceSyntacticConstraintsFeature and TargetPreferencesFeature by Matthias Huck
12. Consistent license of Moses source code to LGPL 2.1, by Philipp Koehn
13. Deterministic cube pruning, by Michael Denkowski
14. Code beautifier, by Jeroen Vermeulen.
15. Phrase-table property, by Matthias Huck

End-to-End Testing and Pre-Made Models

As with previous versions, a number of end-to-end experiments were run. This is the main test to ensure that the Moses pipeline is functional with ‘real-world’ datasets. The translation quality is noted, as measured by BLEU, to ensure that there is no significant variation in performance due to any interaction between components in the pipeline.

We have increased the number of experiments in this release to test the major new functionality, particularly the Implz, Moses2, and the OSM feature. In addition, we also performed a subset of the end-to-end experiments on macOS, Windows 10, Fedora 26, in addition to the out main operating system, Ubuntu 16.04.

The end-to-end tests produce a large number of tuned models. The models, as well as all configuration and data files, are made available for download. This is useful as a template for users setting up their own experimental environment, or for those who just want the models without running the experiments.

The URL for downloading the pre-made models is:

http://www.statmt.org/moses/RELEASE-4.0/models/

The list of experiments and results for each Moses version are shown in Appendix A.

Cross-Platform Compatibility

As with previous releases, we ensured that Moses, IRSTLM, and MGIZA compiles on a number of operating systems commonly used by the community. In this release, on concentrate on 64-bit OSes, and current and future operating system, rather than legacy distributions. We tested them on the following platforms:

i. Windows 10 with Ubuntu 16.04 subsystem
ii. MacOS Sierra with MacPorts
iii. Ubuntu 16.04
iv. Ubuntu 17.04
v. Fedora 26
vi. Centos 7

Binaries for Ubuntu 16.04, Windows 10, and MacOS are made available for download for users who do not wish to compile their own version. The URL for downloads is
http://www.statmt.org/moses/RELEASE-4.0/binaries/
We also make available the 64-bit Linux virtual machines with Moses, IRSTLM, and MGIZA pre-installed at the following URL:
http://www.statmt.org/moses/RELEASE-4.0/vm

There are a number of Issues on some platforms:

1. Regression tests are only guaranteed to pass on Ubuntu 16.04 (64-bit), since different compiler versions and different versions of the standard library can lead to differences in the way floating point numbers are rounded. See, for example http://www.exploringbinary.com/correctly-rounded-conversions-in-gcc-and-glibc/
2. mt-eval13a.pl doesn't run on mac OSX - use mt-eval14.pl.
3. irstlm doesn't compile on ubuntu 17.04 or Fedora 26 . Used KenLM where necessary
4. Using boost 1.65.0 or 1.65.1 causes segfault in Moses2 during regression test

Appendix A - Comparison of BLEU scores with previous versions of Moses

			RELEASE 0.91	RELEASE 1.0	RELEASE 2.1	RELEASE 3.0	RELEASE 4.0
En-es	1	pb	24.70	24.81	24.61	24.66	24.67
	2	hiero	24.18	24.20	23.74	24.45	24.56
	3	(1) + placeholders			24.80	24.59	24.67
	4	(1) + CreateOnDiskPt			24.67	24.74	24.65
Es-en	1	pb	22.87	23.01	22.97	22.97	22.90
	2	hiero	22.32	22.37	22.20	22.26	22.15
	3	(1) + placeholders			22.78	22.65	22.76
	4	(1) + CreateOnDiskPt			23.06	22.83	22.98
En-cs	1	pb	10.98	11.04	11.16	11.03	10.95
	2	hiero	10.92	10.93	10.90	10.80	10.50
	3	(1) + placeholders			10.57	10.67	10.65
	4	(1) + CreateOnDiskPt			11.12	10.88	11.04
	5	(2) + Ken's incre algo			10.76	10.42	10.62
Cs-en	1	pb	16.00	15.72	15.81	15.76	15.77
	2	hiero	15.89	15.68	15.66	15.68	15.53
	3	(1) + placeholders			15.65	15.58	15.64
	4	(1) + CreateOnDiskPt			15.80	15.83	15.75
	5	(2) + Ken's incre algo			15.43	15.70	15.41
En-de	1	pb	11.72	11.87	11.71	11.61	11.71
	2	hiero	11.59	11.62	11.45	11.44	11.44
	3	(1) + CreateOnDiskPt			11.64	11.80	11.62
	4	(1) + POS de	11.55	11.67	11.75	11.71	12.00
	5	(4) + POS en	11.64	11.75	11.69	11.79	11.86
De-en	1	pb	15.71	15.75	15.80	15.75	15.69
	2	hiero	15.74	15.53	15.56	15.91	16.16
	3	(1) + POS de	15.71	15.84	15.70	15.91	15.90
	4	(3) + POS en	15.84	15.93	15.94	15.80	16.11
	5	(1) + CreateOnDiskPt			15.70	16.01	15.75
	6	(3) +			15.76	15.92	16.03

		CreateOnDiskPt					
	7	(4) + CreateOnDiskPt			15.79	15.88	16.13
	8	(1) + lmplz + probing pt					16.05
	9	(2) + lmplz					16.09
	10	(3) + lmplz + probing pt					16.01
	11	(4) + lmplz + probing pt					16.27
	12	(8) + moses2					15.90
	13	(9) + moses2 + probing pt					15.99
	14	(10) + moses2					15.97
	15	(8) + OSM					16.65
	16	(10) + OSM					15.95
	17	(11) + OSM					16.27
	18	(4) + gen					15.98
En-fr	1	pb truecase	24.43	24.38	24.45	24.44	24.40
	2	pb recase	22.95	22.94	22.67	22.64	22.72
	3	(2) + hiero	22.35	22.28	22.11	22.36	22.33
	4	(2) + POS en	23.34	23.05	23.06	22.93	22.92
	5	(2) + kbmira			22.81	23.00	22.88
	6	(2) + pro			22.52	22.37	22.24
	7	(5) + mira			21.59	n/a	22.81
	8	(2) + CreateOnDiskPt			22.89	n/a	
	9	(2) + CompactPt			22.67	22.69	
	10	(1) + lmplz + probing pt					24.80
	11	(2) + lmplz + probing pt					23.18
	12	(3) + lmplz					22.82
	13	(4) + lmplz + probing pt					23.44
	14	(10) + moses2					24.56
	15	(11) + moses2					23.07
	16	(12) + moses2 + probing pt					22.76
	17	(13) + moses2					23.44
	18	(10) + OSM					25.10
	19	(11) + OSM					23.55
	20	(13) + OSM					23.52

Fr-en	1	pb truecase	24.09	24.06	23.87	23.88	24.25
	2	pb recased	22.40	22.41	22.43	22.42	22.38
	3	(2) + hiero	18.19	18.05	17.78	17.94	17.97
	4	(2) + en POS	22.51	22.55	22.47	22.48	22.92
	5	(2) + kbmira			22.46	22.52	22.45
	6	(2) + pro			22.44	22.56	22.47
	7	(2) + CreateOnDiskPt			22.50	n/a	22.43
	8	(2) + Probing pt			22.58	22.50	22.49
	9	(1) + lmplz					24.37
	10	(2) + lmplz					22.67
	11	(3) + lmplz					18.26
	12	(4) + lmplz					23.03
	13	(9) + probing pt					24.37
	14	(10) + probing pt					22.66
	15	(12) + probing pt					23.11
	16	(13) + moses2					24.22
	17	(14) +moses2					22.54
	18	(11) + probing pt + moses2					18.40
	19	(13) + OSM					24.85
	20	(14) + OSM					23.33

Acknowledgements

Thanks to everyone who contributed code and time.

Thanks to the University of Edinburgh for server time.