

Lecture 4: presentation of the SMT lecture

Introduction, agenda, labs,
the “smt machine”

Who am I

- Senior software engineer
 - Computer scientist
 - Interest in languages (fluent: fr,en,it / basic: hi,de / passive: es,br)
 - My skills: IT→NLP→MT
- Between Research and Development
- Work in WIPO: UN agency (intellectual property)
previous jobs: French Medical university, European Commission
- Why MT?
 - WIPO is an early adopter of MT
 - In production: <http://patentscope.wipo.int/transalte/>
 - Recently Neural MT
- Why SMT?
 - We are a team of 2 persons
 - We have to deal with 10+ languages
 - We have a lot of data (e.g. 60 Million sentences Chinese-English)
 - We create specific engines (better than Google Translate on patent texts)

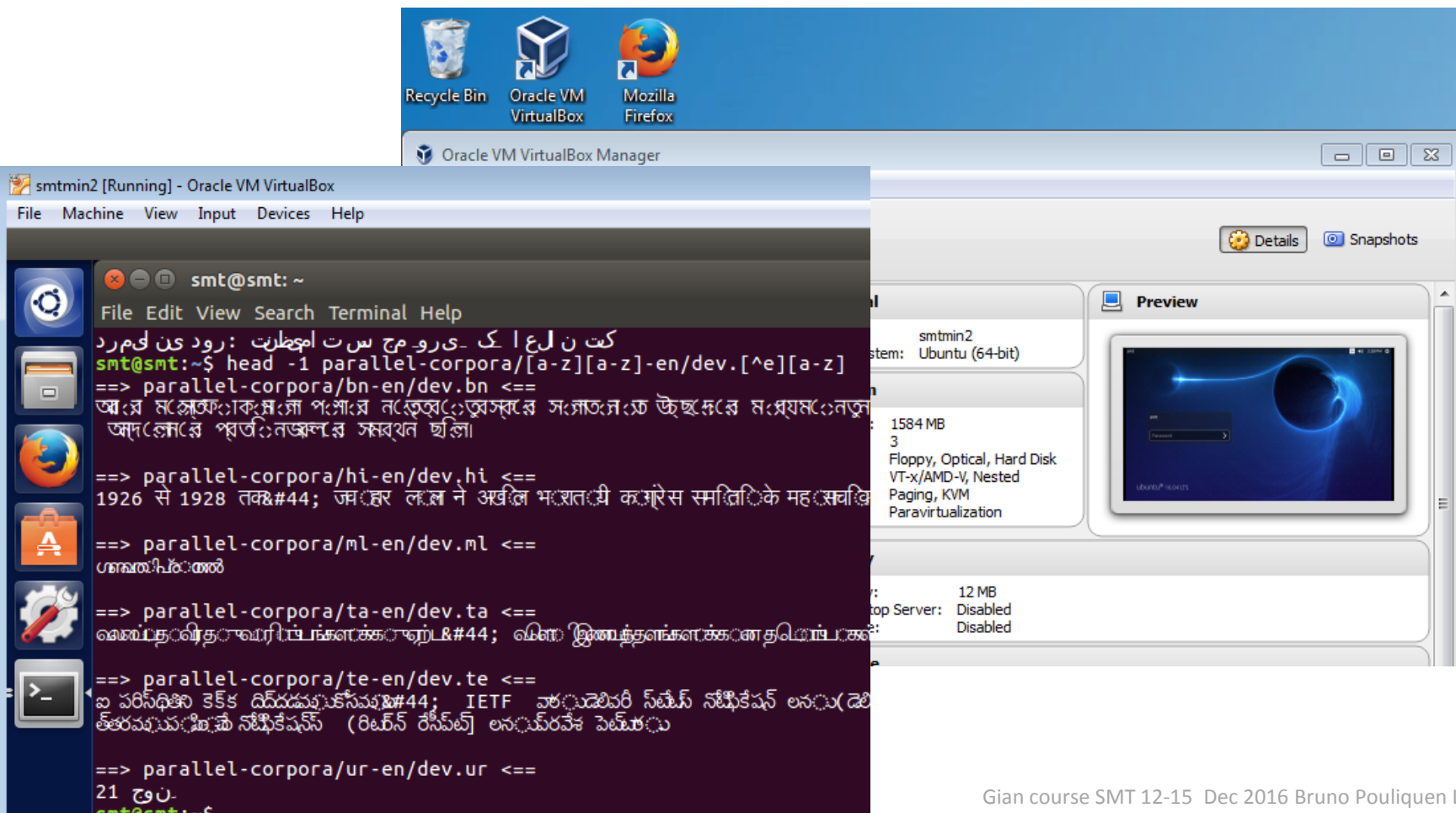
Course agenda

	Lecture	Lab
Monday	SMT	Build your own word-to-word MT (IBM model 1) -or- First steps with Unix commands (browse and test MT)
Tuesday	SMT (deeper)	Installation of a full MT product (Moses) and launch training bn,hi,ml,ta,te,ur→en
Wednesday	Quality	Quality assessment of trained MT models
Thursday	Neural MT	Experiments with simple Neural network machines Usage of NMT models

SMT virtual machine

Install your own virtual machine:

<https://github.com/mlearningbruno/giancourse/blob/master/misc/installSMTtoolsOnVirtualMachineV2.pdf>



More information about this course:

<https://github.com/mlearningbruno/giancourse/>

The screenshot shows the GitHub repository page for `mlearningbruno / giancourse`. At the top, there are buttons for Watch (0), Star (0), and Fork (0). Below this is a navigation bar with links for Code, Issues (0), Pull requests (0), Projects (0), Pulse, and Graphs. The main heading is "Necessary scripts for labs". Below this is a summary bar showing 9 commits, 1 branch, 0 releases, and 1 contributor. There are buttons for "Branch: master", "New pull request", "Find file", and "Clone or download". A commit history table follows, listing recent commits by mlearningbruno. The latest commit is titled "Lectures" and was made 7 minutes ago. Below the table is a section for the README.md file, which contains the title "SMT giancourse" and a description of the repository's purpose.

mlearningbruno / giancourse

Watch 0 Star 0 Fork 0

<> Code Issues 0 Pull requests 0 Projects 0 Pulse Graphs

Necessary scripts for labs

9 commits 1 branch 0 releases 1 contributor

Branch: master New pull request Find file Clone or download

mlearningbruno committed on GitHub Lectures		Latest commit 625ac14 7 minutes ago
labs	Tutorial on how to install virtual machine for labs	4 hours ago
lectures	Lectures	7 minutes ago
misc	create virtual machine	13 minutes ago
README.md	Update README.md	10 minutes ago

README.md

SMT giancourse

SMT Gian course This is the repository where you will find necessary scripts and documentations to run the labs during the GIAN course. The lectures will be available too