上周在Ubuntu16.04上安装了moses这个系统，还好之前有一些Linux的基础，安装过程中遇到的比较大的困难就是一些makefile文件和configure文件需要修改以及出现的各种版本不兼容的问题，好在最终是安装成功了。我把安装过程以及安装过程中遇到的一些问题做了记录，用markdown写在了.md文件。我把内容粘贴在了后面，但是跟word有一些格式上的不兼容，有些图片丢失了。经过一周的Linux使用，现在对Linux熟悉起来了，但是对于Linux下c++通过gcc和boost编译这个过程不太清楚。

现在正在学习如何来使用这个Moses系统，但是由于这个系统的说明文档都是英文的，同时网上的教程比较少，同时又由于在Linux下使用，我看一下Moses系统又是基于c++的，所以学习使用起来有点困难，还需要一点点的摸索。

# MOSES安装记录

###### 环境：Ubuntu16.04

###### 需要安装的工具：giza++,IRSTLM5.80.08,cmph2.0,xmlrpc1.33.14,boost1.64

###### 0 安装依赖的包

sudo apt-get install build-essential git-core pkg-config automake libtool wget zlib1g-dev python-dev libbz2-dev  
sudo apt-get install libsoap-lite-perl

###### 1 从GitHub上clone moses

cd ~  
mkdir smt  
cd smt  
git clone https://github.com/moses-smt/mosesdecoder.git

###### 2 安装giza++

git clone https://github.com/moses-smt/giza-pp.git  
cd giza-pp  
make

把之后需要用到的giza++文件复制到mosedecoder文件夹中创建的tools文件夹下

cd ../mosesdecoer  
mkdir tools  
cp ../giza-pp/GIZA++-v2/GIZA++ ../giza-pp/GIZA++-v2/snt2cooc.out ../giza-pp/mkcls-v2/mkcls tools

###### 3 安装IRSTLM5.80.08

cd ..  
mkdir irstlm  
wget https://jaist.dl.sourceforge.net/project/irstlm/irstlm/irstlm-5.80/irstlm-5.80.08.tgz  
tar zxvf irstlm-5.80.08.tgz  
cd irstlm-5.80.08  
cd trunk  
./regenerate-makefiles.sh  
./configure --prefix=~/smt/irstlm #设置irstlm的安装路径  
make install

###### 4 安装cmph2.0

cd ~/smt  
wget http://www.achrafothman.net/aslsmt/tools/cmph\_2.0.orig.tar.gz  
tar zxvf cmph\_2.0.orig.tar.gz  
cd cmph-2.0/  
./configure  
make  
make install

###### 5 安装xmlrpc1.33.14

cd ~/smt  
wget http://www.achrafothman.net/aslsmt/tools/xmlrpc-c\_1.33.14.orig.tar.gz  
tar zxvf xmlrpc-c\_1.33.14.orig.tar.gz  
cd xmlrpc-c-1.33.14/  
./configure  
make  
make install

###### 6 安装boost1.64

cd ~/smt  
wget https://dl.bintray.com/boostorg/release/1.64.0/source/boost\_1\_64\_0.tar.gz  
tar zxvf boost\_1\_64\_0.tar.gz   
cd boost\_1\_64\_0/  
./bootstrap.sh   
./b2 --layout=system link=static install || echo FAILURE

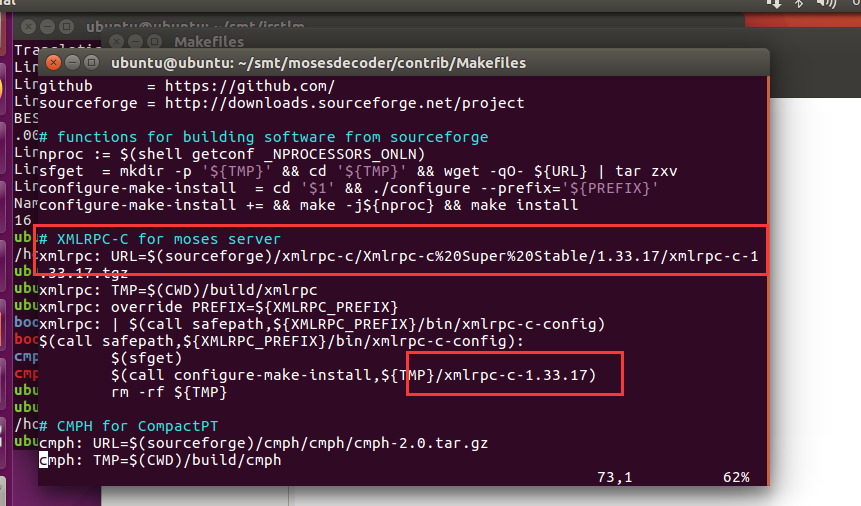
###### 7.1 编译moses

cd ~/smt/mosesdecoder  
make -f contrib/Makefiles/install-dependencies.gmake

在这个过程出现了ERROR

sudo vim install-dependencies.gmake

通过查看这个gmake文件，找到问题的位置。问题应该是出在在这个文件编译过程中xmlrpc这个包下载不下来，看了一些博客和教程他们的方法都不管用，我自己认为xmlrpc之前已经安装过了，所以我注释了这一段。

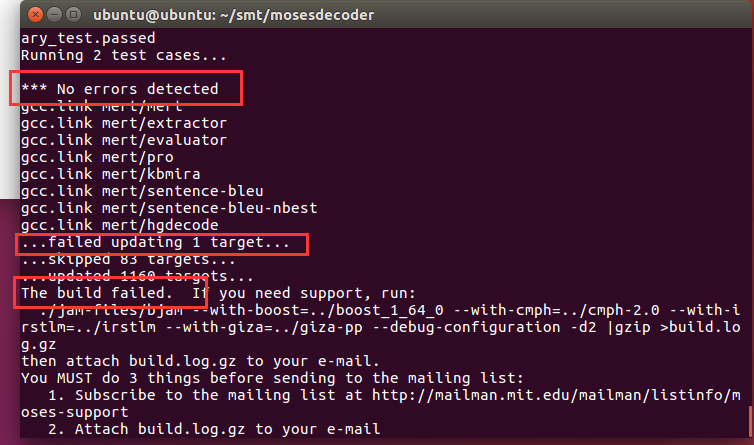


而且在图中可以看到xmlrpc这个包的URL地址，我直接去这个地址上也down不下来，它的版本跟我在前几步中下载安装的版本也不对。

###### 7.2 安装moses

cd ~/smt/mosesdecoder  
sudo ./bjam --with-boost=../boost\_1\_64\_0 --with-cmph=../cmph-2.0 --with-irstlm=../irstlm --with-giza=../giza-pp

这个过程需要一点时间，我在这一步失败了，提示我built failed

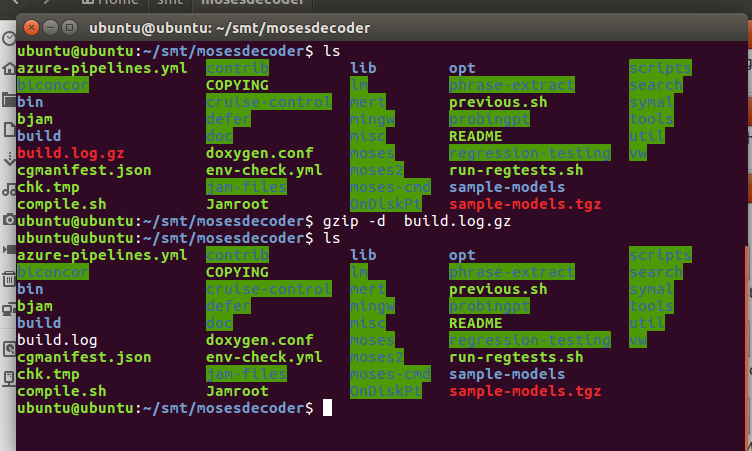


运行如下的命令

sudo ./jam-files/bjam --with-boost=../boost\_1\_64\_0 --with-cmph=../cmph-2.0 --with-irstlm=../irstlm --with-giza=../giza-pp --debug-configuration -d1 |gzip >build.log.gz

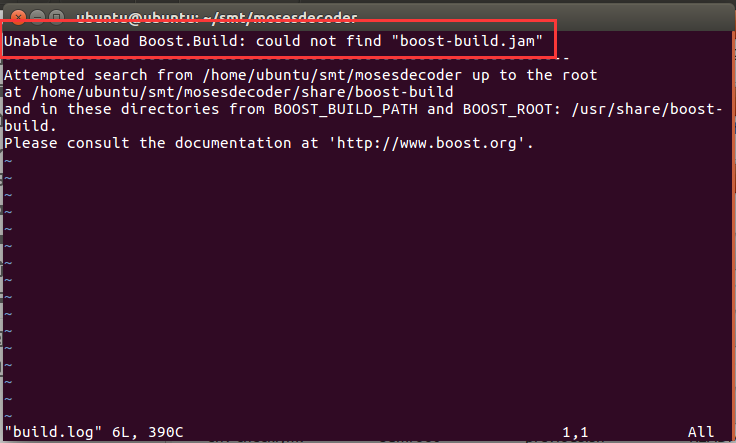
在mosesdecoder文件中找到build.log.gz这个压缩包

# 解压缩  
gzip -d build.log.gz



查看build.log

vim build.log

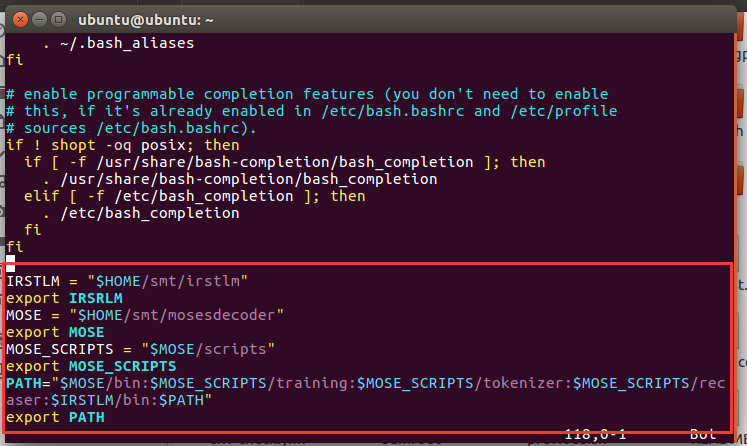


提示找不到boost-build.jam文件，这个问题解释起来比较麻烦，我通过在.bashrc文件中加入路径解决了这个问题。

# .bashrc文件在用户目录下  
cd ~  
vim .bashrc

在文件的最后添加上以下几句

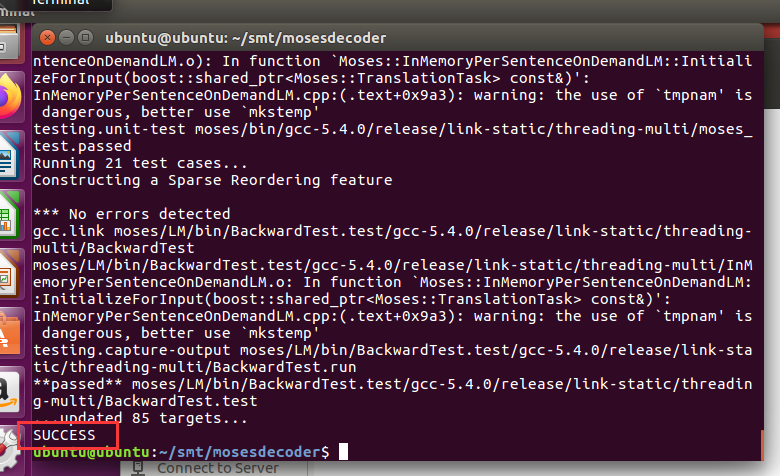
IRSTLM = "$HOME/smt/irstlm"  
export IRSRLM  
MOSE = "$HOME/smt/mosesdecoder"  
export MOSE  
MOSE\_SCRIPTS = "$MOSE/scripts"  
export MOSE\_SCRIPTS  
PATH="$MOSE/bin:$MOSE\_SCRIPTS/training:$MOSE\_SCRIPTS/tokenizer:$MOSE\_SCRIPTS/recaser:$IRSTLM/bin:$PATH"  
export PATH



保存退出后再次进行安装moses

cd ~/smt/mosesdecoder  
sudo ./bjam --with-boost=../boost\_1\_64\_0 --with-cmph=../cmph-2.0 --with-irstlm=../irstlm --with-giza=../giza-pp

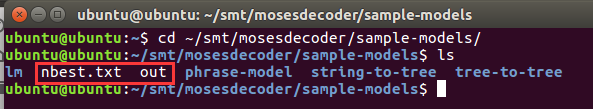
最后出现SUCCESS提示安装成功



###### 7.3 测试moses

cd ~/mosesdecoder  
wget http://www.statmt.org/moses/download/sample-models.tgz  
tar xzf sample-models.tgz  
cd sample-models  
   
# run the decoder  
cd ~/mosesdecoder/sample-models  
~/mosesdecoder/bin/moses -f phrase-model/moses.ini < phrase-model/in > out

在sample-models下面回生成两个文件：nbest.txt和out



可以打开看一下out文件，会显示两行this is a small house，说明安装是成功的

vim out

