

---

Subject: Re: Remove unused languages  
Posted by [crydev](#) on Mon, 06 Jan 2014 15:20:06 GMT  
[View Forum Message](#) <> [Reply to Message](#)

---

I did some basic calculations, based on the information I found in the disassembler and what I found in files: Core.t and LangInfo.cpp.

The image below shows a few example strings. They are located in Core.t and I assume they are translations for an error or information message in multiple languages. However, in the disassembler you can see that they are all compiled into the executable. I did not measure this, but an educated guess on Core.t, which is 44 kb (~ 40 kb with margins of non-trivial information) is linked into the executable. If just the english translations are linked into the executable, the size could be reduced by roughly: average string length divided by the number of languages built into U++.

The image below shows the language array in LangInfo.cpp. The big rectangle shows the strings that are linked, which are a hell of a lot and they shouldn't be necessary.

The small rectangle shows the fact that some strings even indicate only very small differences. The compilation loaded into IDA Pro is built in VC10 Optimal with String Pooling enabled. Even if these strings could not be eliminated, it would be nice if they could be concatenated in order to remove the redundancy that is generated.

If you need any more information, please tell me.

Regards,  
crydev

### File Attachments

---

1) [IDA Strings.PNG](#), downloaded 926 times



2) [Langinfo.PNG](#), downloaded 982 times

.rdata:00551AE8	00000008	C	Bengali\t
.rdata:00551B04	00000012	C	Romansh\tRumantsch\t
.rdata:00551B1C	0000000F	C	Mohawk\tKanien'k
.rdata:00551B34	00000011	C	Irish\tGaeilge\t\b<
.rdata:00551B48	0000002A	C	Inuktitut (Latin)\tInuktitut (Kanatami)\t\b]
.rdata:00551B74	00000007	C	Nepali\t
.rdata:00551B94	00000016	C	Mapudungun\tMapudungun\t
.rdata:00551BB0	0000000E	C	Frisian\tFrysk\t
.rdata:00551BC4	00000007	C	Pashto\t
.rdata:00551BD8	0000000F	C	Luxembourgish\tL
.rdata:00551BE9	0000000D	C	tzebuergesch\t
.rdata:00551BFC	00000012	C	Filipino\tFilipino\t
.rdata:00551C14	00000013	C	Bosnian (Cyrillic)\t
.rdata:00551C3C	00000013	C	Serbian (Cyrillic)\t
.rdata:00551C60	0000001C	C	Sami (Inari)\tsǃmikielǃ\t\$;
.rdata:00551C7C	00000020	C	Sami (Skolt)\tsǃmǃmǃ'Ç©iǃmll\t;
.rdata:00551C9C	00000010	C	Sami (Southern)\t
.rdata:00551CAE	00000012	C	arjelsaemiengiele\t
.rdata:00551CC4	00000010	C	Sami (Southern)\t
.rdata:00551CD6	00000012	C	arjelsaemiengiele\t
.rdata:00551CEC	00000013	C	Sami (Lule)\tjulevus
.rdata:00551D01	00000009	C	megiella\t
.rdata:00551D10	00000013	C	Sami (Lule)\tjulevus
.rdata:00551D25	00000009	C	megiella\t
.rdata:00551D34	00000016	C	Sami (Northern)\tdavvis
.rdata:00551D4C	00000009	C	megiella\t
.rdata:00551D58	00000024	C	Sami (Northern)\tdavvisǃmegiella\t\b;
.rdata:00551D7C	00000016	C	Sami (Northern)\tdavvis
.rdata:00551D94	00000009	C	megiella\t
.rdata:00551DA0	00000020	C	Northern Sotho\tSesotho sa Leboa\t
.rdata:00551DC4	0000000D	C	Zulu\tisiZulu\t
.rdata:00551DD4	0000000F	C	Xhosa\tisiXhosa\t
.rdata:00551DE8	00000010	C	Tswana\tSetswana\t
.rdata:00551DFC	00000011	C	Quechua\ttrunasimi\t
.rdata:00551E10	00000014	C	Quechua\ttrunasimi\t\bk
.rdata:00551E24	00000011	C	Quechua\ttrunasimi\t
.rdata:00551E38	0000000E	C	Maltese\tMalti\t
.rdata:00551E4C	0000000B	C	Maori\tReo M
.rdata:00551E60	0000000E	C	Welsh\tCymraeg\t
.rdata:00551E74	00000017	C	Serbian (Latin)\tsrpski\t
.rdata:00551E90	00000011	C	Bosnian\tbosanski\t
.rdata:00551EA4	00000012	C	Croatian\tthrvatski\t
.rdata:00551EBC	00000014	C	Spanish\tEspañol\tP\n
.rdata:00551ED0	00000014	C	Spanish\tEspañol\tL\n
.rdata:00551EE4	00000014	C	Spanish\tEspañol\tH\n
.rdata:00551EF8	00000014	C	Spanish\tEspañol\tD\n
.rdata:00551F0C	00000014	C	Spanish\tEspañol\t@\n
.rdata:00551F20	00000007	C	Arabic\t