Supplementary Materials

Code-Switched Language Models Using Neural Based Synthetic Data from Parallel Sentences

Dataset Distribution

We provide the dataset distribution of SEAME Phase II as a reference for future work.

Data	Recording list	
	Conversation	**Interview**
Train	02NC03FBX, 02NC04FBY, 03NC05FAX 03NC06FAY, 04NC07FBX, 04NC08FBY 06NC11MAX, 06NC12MAY, 07NC13MBP 07NC14FBQ, 08NC15MBP, 08NC16FBQ 09NC17FBP, 09NC18MBQ, 10NC19MBP 10NC20MBQ, 11NC21FBP, 11NC22MBQ 12NC23FBP, 12NC24FBQ, 13NC25MBP 13NC26MBQ, 14NC27MBP, 14NC28MBQ 16NC31FBP, 16NC32FBQ, 18NC35FBP 18NC36MBQ, 19NC37MBP, 19NC38FBQ 21NC41MBP, 22NC43FBP, 22NC44MBQ 23NC35FBQ, 23NC45MBP, 24NC35FBQ 24NC45MBP, 25NC43FBQ, 25NC47MBP 26NC48FBP, 26NC49FBQ, 27NC47MBQ 27NC50FBP, 28NC51MBP, 28NC52FBQ 29NC53MBP, 29NC54FBQ, 30NC48FBP 30NC49FBQ, 31NC35FBQ, 31NC50XFB 32NC36MBQ, 32NC50FBP, 33NC37MBP 33NC43FBQ, 34NC37MBP, 35NC56MBP 36NC46FBQ, 37NC45MBP, 38NC50FBP 39NC57FBX, 40NC58FAY, 41NC59MAX 42NC60FBQ, 44NC44MBQ, 45NC22MBQ 46NC41MBP	NI02FAX, NI04FBX, NI05MBQ NI06FBP, NI07FBQ, NI08FBP NI09FBP, NI10FBP, NI12MAP NI13MBQ, NI14MBP, NI15FBQ NI16FBP, NI17FBQ, NI18MBP NI19MBQ, NI20MBP, NI21MBQ NI22FBP, NI23FBQ, NI24MBP NI25MBQ, NI26FBP, NI27MBQ NI28MBP, NI29MBP, NI30MBQ NI31FBP, NI32FBQ, NI33MBP NI34FBQ, NI35FBP, NI36MBQ NI37MBP, NI39FBP, NI40FBQ NI41MBP, NI42FBQ, NI43FBP NI44MBQ, NI45FBP, NI46FBQ NI47MBP, NI48FBQ, NI49MBP NI50FBQ, NI51MBP, NI52MBQ NI53FBP, NI54FBQ, NI55FBP NI56MBX, NI57FBQ, NI58FBP NI59FBQ, NI60MBP, NI61FBP NI62MBQ, NI63MBP, NI64FBQ NI65MBP, NI66MBQ, NI67MBQ UI02FAZ, UI03FAZ, UI04FAZ UI05MAZ, UI06MAZ, UI07FAZ UI08MAZ, UI10FAZ, UI11FAZ UI12FAZ, UI13FAZ, UI14MAZ UI15FAZ, UI16MAZ, UI17FAZ UI18MAZ, UI19MAZ, UI20MAZ UI21MAZ, UI22MAZ, UI23FAZ UI24MAZ, UI25FAZ, UI26MAZ UI27FAZ, UI28FAZ, UI29FAZ
Dev	01NC01FBX, 01NC02FBY, 15NC29FBP 15NC30MBQ, 21NC42MBQ, 43NC61FBQ	UI01FAZ, UI09MAZ
Test	05NC09FAX, 05NC10MAY, 17NC33FBP 17NC34FBQ, 20NC39MBP, 20NC40FBQ	NI01MAX, NI03FBX

Table 1: SEAME Phase II