433-460 Project Report

ebe rid n

st No ember 2 4

Intro tion

tivati n

M ne Inf mmati n s a allable t da than e en bef ne. Mllll ns and mllll ns f

The latest question answering conference was NTCIR-4 held in Japan in June 2004 where monolingual Japanese systems were tested. The best systems there were getting MRR (Mean Reciprocal Ranks) of about 0.6 which shows an improvement from past conferences but there is still a long way to before

is shown in Figure 1.

t pe and expe ted answe t pe f the quest n. T e aluate a ainst the ld standad data un ues i n.e al().

nswe T e lassi cati n

This fun tl n takes the t kenlsed questl n and the questl n t pe and attempts t detemmlne an expe ted answe t pe fo mo the Named Entlt t p l .

l enent questl n t pes su est dl enent questl n patte ns and hen e the po essln ls a ndln t questl n t pe. l n mo st questl n t pes the enenal mle patte n ls t l k at the w nds just bef ne o just afte the lnte n atl e ln the questl n sea hln f o a sallent ente w nd. lf this ente w nd ls f und ln the named entlt t p l lt ls etumed as the expe ted answe t pe. A l kup table was nstmu ted t tmansf mmo s me mmo n abble latl ns and s n n ms t thelo equl alent ln the NE t pl su h as n entln 名 (po nam) t レビ 名 (TVepo nam) o 社名 (mpan np natl n) t 企 名 (np - natl n).

lf n nele ant ente w nd ls f und a mo ne enenal entlt that mat hes the questl n t pe ls etumed. l n example the questl n t pe どこ (whene) lf n mo ne spe lfi lnf mmatl n ls f und w uld etum an answe t pe f 名 (l atl n).

lt

lt was h ped that mo ne questl n data mul ht be a allable t e aluate the s stem a alnst but that has n t e entuated at this tlme. The data a allable was used t nstmu t enenallsed mules and then see h w l sel these mules uld po du e the same questl n and answe t pes l en ln the nl lnal data.

lt was f und that the questl n t pe uld be dete ted wlth an a una f an und 98 a alnst the nl lnal data. Examlnln the ln nne tl lasslfied questl ns lt was f und that n s me asl ns the lasslfi atl n f the " ld standand data was ln nslstent and f o example lasslfied questl ns that wene leanl askln "h w man o "h w mu h as slmpl "what . hllst the po - nam uld ha e been mo dlfied t lasslf these questl ns a ndln t the ld standand l en that the pump se f thls s stem ls t anal sls Japanese questl n ln enenal as pp sed t these questl ns ln pantl ulan these mo dlfi atl ns wene n t made.

The then em ns f und ln questl n lasslfi atl n wene ln enenal due t ln o- ne t se mentatl n b Cha en whl h led t lntem atl es n t beln ldentlfied. Vanl us patte ns uld ha e been hand- ded .944-l- pa nam thel the te mentatl n exo ns ean t nslstent n ndk-o und o these ls

A ENDIX

A Q tion Typ

\ \ h \ w an h \ w mu h when h \ w where