A Examples of HAN Attention

The examples were taken from the Spanish-English TED talks corpus. We show the behavior of the attention function of HAN. First, we show the attention to context for HAN encoder and HAN decoder respectively. Second, we show the multi-head attention only for HAN decoder (English) for better understanding.

A.1 Encoder and Decoder Attention

Currently Translated Sentence

```
Src.: y toqué el primer movimiento del concierto para violín de Beethoven .

Ref.: and I played the first movement of the Beethoven Violin Concerto .
```

Base: and I *touched* the first move from the concert to Beethoven.

Cache: and I *touched* the first move of Beethoven 's violin . HAN: and I **played** the first move of Beethoven 's violin .

Context from Previous Sentences

HAN decoder context with target. *Query*: played (En) and he was talking about invisible demons and smoke , and how someone was sleeping with him . and I felt fear , not for me , but fear that I was going to lose it ... so I just started playing . HAN encoder context with source. *Query*: toqué (Es) y hablaba de demonios invisibles y humo , y de cómo alguien lo estaba envenenando mientras dormía . y yo sentí miedo , no por mí , sino miedo de que iba a perderlo ... so por ello sólo empecé a tocar .

Table 1: In this example, the HAN model disambiguates correctly the word "toque", which can be translated as "touched" or "played". We can see that the HAN decoder uses the semantically close word "playing" from the previous sentence. In similar manner, the HAN encoder focused on "tocar" which is coherent with "toque".

A.2 Multi-Head Attention

Currently Translated Sentence

```
Src.: y < ellos > estarían tan compenetrados en la partida de dados porque los juegos son tan atractivos ...
Ref: and they would be so immersed in playing the dice games because games are so engaging ..
Base: and you would be so < unk > in the start of it because games are so attractive ...
HAN: and they would be so < unk > in the start of dice
```

Context from Previous Sentences. Query: they

because games are so attractive ...

```
Head 2: Attention to the antecedent "people" in s^{t-3}.
   people suffered . people suffered .
   it was an extreme situation . they needed an extreme s
   so, according to Indyk, the games of dice and a polic
   y was established throughout the kingdom : one day , e
   verybody would eat , and the next day , everybody woul
Head 4: Attention to the same pronoun "they" in s^{t-2}
   people suffered . people suffered .
   it was an extreme situation . they needed an extreme s
   so , according to Indyk , the games of dice and a polic
   y was established throughout the kingdom : one day , e
   verybody would eat , and the next day , everybody woul
Head 7: Attention to verbs that conjugate with "they"
   people suffered . people suffered
   it was an extreme situation . they needed an extreme s
   olution .
    so , according to Indyk , the games of dice and a polic
   y was established throughout the kingdom : one day , e
   verybody would eat, and the next day, everybody woul
```

Table 2: This example displays the translation of Spanish pronoun "ellos", which is a dropped-pronoun which is implicit in the verb conjugation of "estarían". As we can observe, HAN translates correctly the dropped-pronoun into the English "they". Each head focuses on a different aspect during translation, for example head 2 seems to attend to the antecedent of the pronoun "people" in the third previous sentence, head 4 attends to the same pronoun on the second previous sentence, and head 7 attends to different verbs on all previous sentences.

Currently Translated Sentence

Src: y como resultado **construimos** relaciones sociales más fuertes .

Ref: and we actually **build** stronger social relationships as a result .

Base: and as a result, we *construct* stronger social relationships.

HAN: and as a result , we **build** stronger social relationships .

Context from Previous Sentences. Query: build



Table 3: This example displays the translation of the ambiguous Spanish word "construimos", which can be translated as "construct" or "build". HAN translates this word correctly according to the context using for example related words "trust", "ties", and "cooperation" on previous sentences with head 1, and a previous translation "builds" in the previous sentence with head 4.

Currently Translated Sentence

Src: antes de los fantásticos controladores de **juegos** teníamos tabas de oveja .

Ref: before we had awesome **game** controllers , we had sheep 's knuckles .

Base: before the fantastic *TV* controllers , we had < unk > .

HAN: before the fantastic **game** controllers , we had <

Context from Previous Sentences. Query: game

unk > .

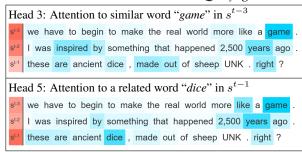


Table 4: This example shows the translation of the Spanish word "*juegos*". The baseline translates it incorrectly, while HAN translates it correctly by spotting a similar translation "*game*" in the third previous sentence with *head 3*, and a related word "*dice*" on previous sentence with *head 5*.