

Machine Translation HW1

Arijit Nukala

September 8, 2023

1 Web Pages Translated

For this assignment, I used Sir Authur Conan Doyle's "The Hound of the Baskervilles" available on bilinguis.com. The website allows for a side by side view of the book in translated languages and English. For this assignment, I will be comparing the Italian translation of the book, using the English translation as the "Human Correction of the Machine Translation" and Google Translate as my Machine Translator. The analysis of the article will be formatted as:

- (a) Spanish Sentence
- (b) Machine Translation
- (c) Human Correction of the Machine Translation
- (d) Assessment of Error

2 Analysis

1. Sentence 1

- (a) Scommetto che lei ha gli occhi anche sulla nuca!
- (b) I bet she has eyes on the back of her head too!
- (c) I believe you have eyes in the back of your head.
- (d) In this translated sentence, Watson is speaking to Sherlock, a male, and the machine translation incorrectly identifies the speaker as a female. Additionally, the translation speaks of Holmes in the third person instead of referring to him as "you".

2. Sentence 2

- (a) Vediamo se lei riesce a ricostruirmi l'uomo dall'esame del suo bastone.
- (b) Let's see if you can reconstruct the man for me by examining her cane.
- (c) Let me hear you reconstruct the man by an examination of it.
- (d) Here, "her cane" in the translation is incorrect because the cane belongs to a man. There are some other minor errors such as the difference in tense between "examining" and "examination". However, the sentence still reads similarly with these other errors.

3. Sentence 3

- (a) - Io penso - dissi, seguendo per quanto mi era possibile i metodi del mio amico - che il dottor Mortimer deve essere un medico in età, con una buona clientela, e molto stimato, dal momento che coloro che lo conoscono gli hanno offerto questo pegno della loro ammirazione.
- (b) "I think," said I, following as far as I could the methods of my friend, "that Dr. Mortimer must be an old doctor, with a good clientele, and highly esteemed, since those who know him have offered him this token. of their admiration.
- (c) "I think," said I, following as far as I could the methods of my companion, "that Dr. Mortimer is a successful, elderly medical man, well-esteemed since those who know him give him this mark of their appreciation."
- (d) The machine translation includes a misplaced period after token. Google translate also uses "old doctor, with a good clientele" in place of "a successful, elderly medical man", which has a different tone and omits the direct explanation of the doctor's status as successful.

4. Sentence 4

- (a) Holmes non mi aveva mai lodato tanto in vita sua e devo ammettere che le sue parole in quel momento mi procurarono una viva soddisfazione, poiché spesso ero rimasto ferito dalla sua indifferenza verso l'ammirazione che io avevo sempre manifestata nei suoi confronti e verso i miei tentativi di dare pubblicità ai suoi metodi.
- (b) Holmes had never praised me so much in his life and I must admit that his words at that moment gave me great satisfaction, since I had often been hurt by his indifference towards the admiration that I had always shown towards him and my friends. attempts to publicize his methods.
- (c) He had never said as much before, and I must admit that his words gave me keen pleasure, for I had often been piqued by his indifference to my admiration and to the attempts which I had made to give publicity to his methods.
- (d) The translation makes an error after "my friends" by including an unnecessary period. "Piqued" does not exactly match "hurt" in the machine translation in terms of tone.

5. Sentence 5

- (a) - Temo, mio caro Watson, che la maggior parte delle sue conclusioni siano sbagliate.
- (b) -I fear, my dear Watson, that most of his conclusions are wrong.
- (c) "I am afraid, my dear Watson, that most of your conclusions were erroneous.
- (d) The translation includes the dash in the Italian grammar and incorrectly uses "his conclusions" instead of "your conclusions". "Your conclusions" is correct because in this sentence, Holmes is speaking directly to Watson, which the machine translator does not understand.

6. Sentence 6

- (a) - Ma poi basta.

- (b) - But then that's enough.
- (c) "But that was all."
- (d) The translation does not convey the meaning of "but that was all." The translation gives a feeling of completion, since "that's enough" implies there should be no more. However, The actual meaning of "that was all" is that there was more that could be said.

7. Sentence 7

- (a) Scoppiai in una risata incredula, mentre Sherlock Holmes si sprofondava nel suo divanetto lanciando verso il soffitto anelli di fumo indefiniti.
- (b) I burst out laughing in disbelief, while Sherlock Holmes sank into his sofa, throwing vague rings of smoke towards the ceiling.
- (c) I laughed incredulously as Sherlock Holmes leaned back in his settee and blew little wavering rings of smoke up to the ceiling.
- (d) The translation mixes "vague" and "wavering" in this example. "Wavering" clouds of smoke slowly fade away while "vague" clouds make one unsure what the clouds are.

8. Sentence 8

- (a) - Per quel che riguarda l'ultima parte non ho dati sufficienti per controllare le sue asserzioni - dissi - ma comunque non è difficile stabilire alcuni particolari inerenti all'età del nostro uomo e alla sua carriera professionale.
- (b) -As regards the last part, I don't have enough data to verify his assertions - I said - but in any case it is not difficult to establish some details relating to our man's age and his professional career.
- (c) "As to the latter part, I have no means of checking you," said I, "but at least it is not difficult to find out a few particulars about the man's age and professional career."
- (d) In the translated sentence, Watson says "I don't have enough data to verify HIS assertions" to Holmes, while speaking directly to Holmes. The translation is incorrect and Watson should be using "you".

9. Sentence 9

- (a) Autore di Capricci dell'atavismo (Lancet, 1882), Siamo in progresso? (Journal of Psychology, marzo 1883).
- (b) Author of Caprices of Atavism (Lancet, 1882), Are we in progress? (Journal of Psychology, March 1883).
- (c) Author of 'Some Freaks of Atavism' (Lancet 1882). 'Do We Progress?' (Journal of Psychology, March, 1883).
- (d) Here, "Caprices" (unpredictable changes in mood) and "Freaks" do not have the same meaning though are similar. Additionally, "are we in progress?" and "do we progress" have different verb tenses.

10. Sentence 10

- (a) So per esperienza che a questo mondo soltanto un uomo simpatico riceve attestati di amicizia, soltanto un uomo privo di ambizioni abbandona una professione a Londra per ritirarsi in campagna, infine soltanto un uomo distratto lascia il proprio bastone anziché il proprio biglietto da visita dopo averti aspettato per un'ora nella tua stanza.
- (b) I know from experience that in this world only a nice man receives certificates of friendship, only a man without ambition abandons a profession in London to retire to the countryside, finally only a distracted man leaves his cane instead of his business card after waiting for you for an hour in your room.
- (c) It is my experience that it is only an amiable man in this world who receives testimonials, only an unambitious one who abandons a London career for the country, and only an absent-minded one who leaves his stick and not his visiting-card after waiting an hour in your room."
- (d) The last part of the list in the translated sentence, "finally only a...", is not the correct method of ending a comma separated list. Instead, "and" should be used. The machine translation is grammatically incorrect.

11. Sentence 11

- (a) - E il cane?
- (b) - It's the dog?
- (c) "And the dog?"
- (d) "It's" and "And" give "the dog" two different meanings. The machine translation refers directly to the dog, questioning if "the dog is" something. The actual translation adds the dog in addition to a question previously asked.

12. Sentence 12

- (a) - Esatto.
- (b) - Exact.
- (c) "Yes, sir."
- (d) "Exact" and "Yes, sir" have differently meaning. The machine translation could be "Exactly", which could be closer to correct because "Exactly" is similar to "yes". However, the tense of "Exact" does not match that of "Exactly" or "yes".

13. Sentence 13

- (a) Il dottor Mortimer ammiccò da dietro le lenti con un'espressione di mite stupore.
- (b) Doctor Mortimer blinked from behind his glasses with an expression of mild amazement.
- (c) Dr. Mortimer blinked through his glasses in mild astonishment.
- (d) The words "amazement" and "astonishment" are similar but do not convey the same meaning in speech. Amazement can be positive while astonishment is almost always in shock negatively.

14. Sentence 14

- (a) - Lietissimo di conoscerla: ho inteso fare il suo nome unitamente a quello del suo amico.

- (b) -Delighted to meet you: I intended to mention his name together with that of his friend.
- (c) "Glad to meet you, sir. I have heard your name mentioned in connection with that of your friend.
- (d) "I intended to mention his name" makes it seem like the subject hoped to speak his name aloud whereas the correct translation of "I have heard your name mentioned" means that the subject has heard his name and does not hope to speak it. The meanings of these sentences are different so the translation is flawed.

15. Sentence 15

- (a) Ora, poiché ammetto che lei in materia è il secondo esperto d'Europa...
- (b) Now, since I admit that you are the second expert in Europe on the subject...
- (c) Recognizing, as I do, that you are the second highest expert in Europe ——"
- (d) The translation "second expert" and "second highest expert" have two different meanings. "Second highest expert" conveys that the person is the second best in the subject whereas "second expert" is more vague.

16. Sentence 16

- (a) - Durante tutto il tempo in cui ella ha conversato con me, ha offerto ai miei occhi circa un paio di pollici del manoscritto;
- (b) -During the whole time that she conversed with me, she offered to my eyes about a couple of inches of the manuscript;
- (c) "You have presented an inch or two of it to my examination all the time that you have been talking.
- (d) While the translator is not blatantly wrong, the translated sentence reads awkwardly. The translation "She offered to my about about a couple of inches of the manuscript" does not sound correct because it sounds like "she" is offering the manuscript to the subject's eyes and not to the subject himself. Therefore, the translation is an error.

17. Sentence 17

- (a) Abbiamo frequentemente registrato in queste colonne le sue generose donazioni ai locali enti di beneficenza e alla contea.
- (b) We have frequently recorded in these columns his generous donations to local charities and to the county.
- (c) His generous donations to local and county charities have been frequently chronicled in these columns.
- (d) The machine translated "Donations to local charities and to the county" makes it sound as if the donations were to both the local charities and to the county whereas the correct translation of "donations to local and county charities" specifies that the donations were made only to charities. The machine translation is incorrect in this case.

18. Sentence 18

- (a) - E non l'ha detto a nessuno?

- (b) -And she didn't tell anyone?
- (c) "And you said nothing?"
- (d) Instead of referring to the subject as "you", the machine translation was not able to pick apart context and referred to the subject as "she".

19. Sentence 19

- (a) - Però non si era avvicinata al corpo?
- (b) -But she hadn't approached her body?
- (c) "But it had not approached the body?"
- (d) Similarly, the machine translation is unable to understand context and refers to the beast as "she" instead of "it".

20. Sentence 20

- (a) - Com'era la notte?
- (b) - How was the night?
- (c) "What sort of night was it?"
- (d) "How was the night" and "What sort of night was it" are two different questions in translation. The machine translation seems to be asking about the quality of the night whereas the actual translation asks about the type of night, referring to weather. While the wording is similar, subtle differences in meaning make the machine translation incorrect.

3 Summary

Based on all of the incorrect translations analyzed, a few common issues between Italian to English machine translations became obvious when analyzing this specific text. However, the ground truth Italian translation of the book did not exactly match the ground truth English version due to idiomatic and structural differences. When parsing the corpus, I attempted to choose sentences where the translation seemed closer to direct and tried to ignore minor differences in wording or structure that did not affect the meaning of the sentence broadly.

One of the major issues is "hallucination". There were situations in which periods appeared after seemingly random words in a sentence. This should not occur at all and I found it puzzling to find this issue. Another major issue is that possessive pronouns were often mismatched. For instance, "he" and "she" were used when "you" should have been used. This could be because in Italian nouns are gendered. The machine translation should take this into account because otherwise, especially in direct speech, the translation can make speakers sound as if they are talking to subjects in the third person. Most other issues were more case specific and involved word order and word translation. For example, the translation of "Exact" to "Yes, sir" does not work because "Exact" should be in the adverb tense "Exactly" for the usage to make sense. Scenarios such as comma separated list termination, Italian quotation notation to English quotation notation conversion, and idiom translation are not

captured by the machine translation system.

In this assignment, I have only chose examples in which machine translation fails. Machine translation systems do an incredibly good job of converting languages between each other an despite shortcomings, have increasingly become a viable way to translate languages accurately. However, as the assignment displays, humans can still decipher real translations from machine translations and therefore, the systems still can benefit from progress.