Delta University Scientific Journal
Journal home page: https://dusj.journals.ekb.eg

The Future of Human Translation in the Artificial Intelligence Era

Author name [Dr. Dina Gaber Zanaty]

Author affiliations [Assistant Professor of Applied Linguistics, Delta University for Science and Technology, Faculty of Arts, English Language and Literature Department]

Correspondence: [Dina Gaber Zanaty.]; [Gamasa, 11152, Egypt.]; Tel [+201005168752]
Email :Dina.Zanaty@deltauniv.edu.eg

ABSTRACT

Advancements in Artificial Intelligence (AI) and Deep Learning (DL) have greatly enhanced the accuracy and quality of Machine Translation (MT). Some argue that human translation is no longer necessary. Human faults are often addressed by self-invented solutions. The development of neural networks in machine translation has led to claims that intelligent systems can now translate as well as humans. Despite advancements in AI, language processing and translation remain challenges. Designing intelligent systems involves inherent biases. The way we design systems is influenced by our personal experiences and biases. The methodology followed in this research which is based on (Hunt at al., 2017) framework is language processing utilizing primary contextual and semantic analysis with reference to comprehensive dictionaries (formed by integrating dictionaries, thesauri, and databases of language and jargon awareness), along with connotation and contextual connotation databases, to perform a thorough parsing of text into parts of speech. The diversity of language structures and cultures makes it unlikely that intelligent robots, even with deep learning skills, can handle them as efficiently as humans can.

Keywords: *Translation, AI, Deep Learning, NLP, Machine Translation, Neural Machine Translation.*

1. Introduction

Advancements in Artificial Intelligence (AI) and Natural Language Processing (NLP) have greatly enhanced the accuracy and quality of Machine Translation (MT) across languages. Some argue that human translation is no longer necessary. After all, human faults are consistently addressed by its own innovations. AI-powered software can now match human translation skills, according to recent claims.

Despite advancements in AI, language processing and translation remain challenging. In late 2016, Google constructed a neural network to enhance its translation algorithms (Pring-Mill, 2018). However, there are several difficulties with the translations, particularly with the social and grammatical aspects of the language.

Machine Translation (MT) now recognizes subtleties in utterances, rather than just syllables, words, phrases, or sentences. The problem of contextualization remains present. To better understand the potential of AI in translation, it's important to examine the limitations that now limit its application.

The semantics of a language, including words, phrases, sentences, and paragraphs, is influenced by cultural context. Meaning can vary depending on the speaker's attitude, goals, and context. Machine translation struggles with idioms, sarcasm, irony, humor, and other literary elements.

Google's translation highlights the limitations of machine translation. Translations may be fragile, particularly when dealing with concepts and emotions. While potentially valuable, technology cannot replace human translations.

Machine translation struggles with language-specific difficulties, as well as social and personal characteristics like age and gender.

AI-powered natural language processing struggles to manage linguistic complexity. The issues vary from syntactic to semantic, as the input language item may not have a Machine translation that aims to accurately represent the target language. Until AI can decipher the meaning of an input language item and encode it into the destination language, it is unlikely that machine translations will replace human translations.

Inherent biases in the design of intelligent systems arise from their status as human creations. The manner in which these systems are designed reflects the perspectives and social experiences of their creators, thereby embedding a biased worldview into their functionality.

Although neural networks have significantly enhanced the quality and accuracy of machine translation, persistent issues and inefficiencies remain, which will be explored further in this paper. Given the diverse linguistic structures and cultural contexts they must manage, achieving human-level efficiency in translation through intelligent machines, even with advanced deep learning capabilities, appears highly unlikely at present.

2. Literature Review:

2.1. AI Translation: Maturation and Advancements

Machine translation (MT), a subfield of computational linguistics, has evolved alongside advancements in computer technology, information theory, and linguistics. Initially based on dictionary-like approaches, MT has progressed through corpus-based computer-aided translation (CAT) to the current state of artificial intelligence-driven neural machine translation (NMT), (Zong, 2018).

Although the foundations of machine learning (ML) can be traced back to the 1990s, the extensive availability of spoken and written data—facilitated by organizations such as the Linguistic Data Consortium (LDC)—has accelerated its development in the 21st century. Notable datasets such as the Penn Treebank (Marcus et al., 1993), Prague Dependency Treebank (Hajic, 1998), PropBank (Palmer et al., 2005), Penn Discourse Treebank, and TimeBank (Pustejovsky et al., 2003) have provided rich linguistic resources with diverse syntactic, semantic, and pragmatic annotations. These resources have spurred further research into parsing, semantic analysis, and other areas such as word sense disambiguation (Palmer et al., 2001; Kilgarriff and Palmer, 2000), question answering, and summarization. Consequently, the development of techniques such as support vector machines, maximum entropy models, multinomial logistic regression (Berger et al., 1996), and graphical Bayesian models has advanced the field of machine translation.

Advancements in high-speed computing have facilitated the application of unsupervised statistical methods and the progress in statistical approaches and topic modeling (Blei et al., 2003) to scenarios where only unannotated data is available. This shift has enabled machine translation systems to operate without the necessity of a reliably annotated corpus.

The latest developments in machine translation involve the deployment of neural networks, which efficiently handle large volumes of linguistic data, thereby significantly enhancing the accuracy of (AI Translation). Major technology companies, such as Google, Facebook, Amazon, and Baidu, have incorporated NMT (AI Translation) into their platforms, enabling real-time multilingual functionalities, including simultaneous translation.

End-to-end Neural Machine Translation (Bahdanau et al., 2015; Cho et al., 2014; Sutskever et al., 2014) has rapidly gained prominence in the field. Although these systems are primarily designed for handling single language pairs, the architecture can be adapted to support multiple language pairs without substantial modifications. By introducing an additional "artificial token to the input sequence to indicate the target language" (Johnson et al., 2017), while maintaining the standard components of the AI Translation model—namely, the encoder, decoder, attention mechanism, and shared word piece vocabulary (Wu et al., 2016)—this model can effectively manage multiple languages. Deep Learning (DL) technologies enable these models to process extensive datasets across multiple layers, allowing for various levels of data abstraction (LeCun et al., 2015).

3. Methodology

The methodology followed in this research is the based on the framework of (Hunt et al., 2017) to Resolve the Meaning of a Body of Natural Language Text Using Artificial Intelligence Analysis in Combination with Semantic and Contextual Analysis (Hunt et al., 2017) it is described as a method of language processing utilizing primary contextual and semantic analysis with reference to comprehensive dictionaries (formed by integrating dictionaries, thesauri, and databases of language and jargon awareness), along with connotation and contextual connotation databases, to perform a thorough parsing of text into parts of speech. A secondary artificial intelligence analysis module employs the output of the primary analysis to adjust parameters and values within the AI module. This secondary module processes iteratively until ambiguities are resolved.

Following the completion of primary and secondary analyses, a ranking matrix processor module evaluates the information gathered by these prior modules to produce a ranking matrix. This matrix represents the semantic content of the text in a format that can be easily utilized by machines or external parties to interpret the text's meaning. Specialized comprehensive dictionaries can be developed for this methodology to achieve particular objectives, such as facilitating cross-language translations or comparing translations across different languages to identify inconsistencies.

AI Translation represents a cutting-edge approach in Artificial Intelligence (AI), predicated on a system that can be trained to discern patterns within data. This system transforms the input data, i.e., the source language to be translated, into the desired output, i.e., the target language or the language into which the translation is performed (Zhou, 2018).

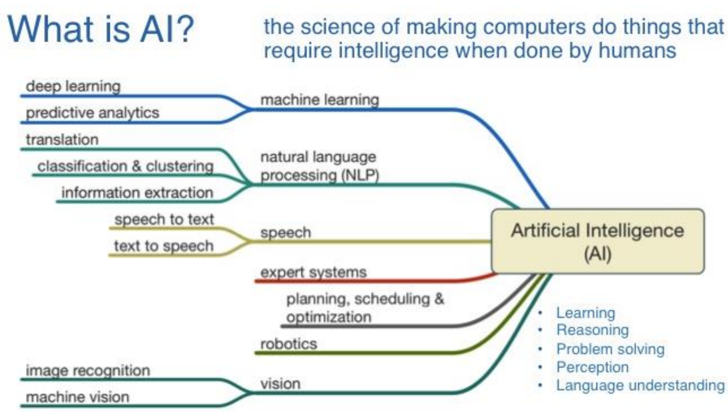

Figure 1. Artificial Intelligence (AI), and Deep Learning utilizing neural networks (adapted from Grasso, as cited in Zhou, 2018).

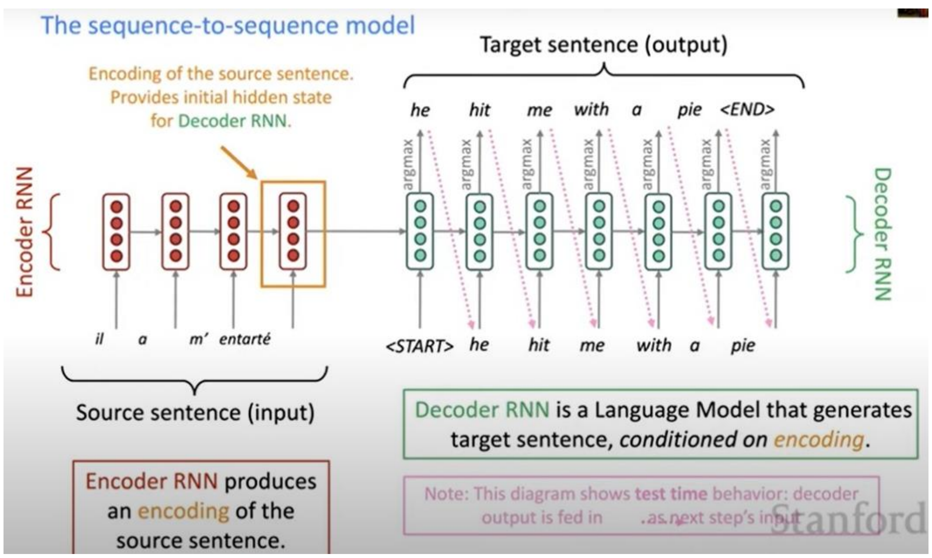

Figure 2 depicts Neural Machine Translation in Google Translate, based on Abigail See's presentations for Stanford's CS224n class, as quoted in Zhou (2018).

The example above illustrates the translation process from French to English. In this process, a French sentence is input into a neural network, with each word represented by numerical vectors. These vectors undergo a series of mathematical transformations, resulting in a sequence of numbers that represent the output sentence in English (Zhou, 2018).

As noted by Hong (2018), "the construction of a corpus is a critical step for achieving successful machine translation, particularly from the perspectives of information mining, retrieval, and processing. Hong further explains that "the retrieval system employs web crawlers to gather network information and uses automatic tagging technology to index this information. Language processing techniques are then applied to establish correspondence between languages and create an indexed database. In the era of artificial intelligence, machines can track users' searches and queries to record, extract, and provide feedback on various translations, thereby contributing to the development of a new corpus. This approach enhances the scope and accuracy of machine translation, aiming to automate the laborious aspects of human translation, improve speed, and reduce costs (Hong, 2018).

In addition to possessing a substantial and comprehensive corpus, (AI Translation) systems are evaluated by translating new sentences that were not part of their initial training. This process aids in refining the system's ability to generalize data that was not previously encountered.

For modern European languages such as English, French, Spanish, and German, there is an abundance of 'translated' data available on the web. This wealth of data facilitates the training of (AI Translation) systems, allowing them to handle diverse syntax, etymology, style, and context effectively. Consequently, neural networks can be better trained on these languages, particularly when dealing with unfamiliar linguistic structures due to their exposure to various patterns.

Conversely, languages that are less prevalent on the web, such as Hindi, Urdu, Greek, Latin, Malay, and Georgian, or their combinations, present a challenge. There is often insufficient data available online to train deep neural networks effectively for these languages. As a result, machine translations are generally more accurate for languages with abundant web data. To achieve similar translation quality for lesser-known or regional languages, it is necessary to increase the volume of data available for training neural network systems.

4. Analysis and Discussion

4.1 Simultaneous Translation and AI Translation

The complexity of simultaneous translation is exacerbated by the variation in word order across different languages. For example, translators working from German to English at the United Nations often face delays due to the late appearance of the verb in German sentences. Contemporary machine translation systems have addressed this challenge through anticipation techniques that facilitate more effective simultaneous translation (Hao, 2018).

Central to any AI-driven machine learning approach to translation is the dataset used for training. Consequently, translation accuracy varies according to the structural similarities between languages. The likelihood of accurate translation increases when the source and target languages exhibit similar sentence structures, whereas dissimilar languages tend to result in higher rates of inaccuracy.

Liang Huang, Chief Scientist at Baidu Research, as cited in Hao (2018), acknowledges the significant challenges faced by human interpreters during simultaneous translation: "Simultaneous translating for human interpreters is extremely challenging and burdensome, so we’re hopeful machines can step in and really make this service more accessible for professionals and consumers."

Although advancements in neural machine translation are promising, the technology remains inferior to human translation in several respects. The results often exhibit shortcomings and fail to capture the nuanced contextual subtleties that human translators manage. Machines currently lack the ability to replicate human perceptions, sensibilities, common sense, reliability, and, notably, memory. Consequently, there are numerous issues with machine translation, as illustrated by various examples from Google Translate output between English and Arabic. The choice of this language pair, aside from the author's proficiency in both languages, underscores the research focus on the challenges associated with translating less commonly represented languages on the web. These languages not only differ in orthography and basic syntax, such as word order, but also exhibit significant cultural distinctions.

4.2. Deficiencies in Cultural Equivalence and Knowledge

Artificial intelligence translation, when applied to certain common English expressions, often struggles to accurately convey the intended meaning in the target language, such as Arabic. This issue is indicative of a broader problem with (AI Translation), as evidenced by the following examples of English-to-Arabic translations, which often result in significant misunderstandings of the intended meaning.

Here are some examples of how (AI Translation) systems can struggle with cultural idioms when translating from English to Arabic, along with a description of how these issues might manifest in translations.

Example 1: "Break the ice"

English Idiom: "Break the ice"

Intended Meaning: To initiate conversation in a social setting to ease tension.

Literal Arabic Translation: "كسر الجليد" (Kasar al-jaleed)

Issue: The literal translation "كسر الجليد" means "breaking the ice," which might be confusing or meaningless in Arabic as it does not convey the idiomatic meaning of easing social tension. The cultural idiom for this concept in Arabic might be different, and the literal translation fails to convey the appropriate context.

Example 2: "Bite the bullet"

English Idiom: "Bite the bullet"

Intended Meaning: To endure a painful or difficult situation with courage.

Literal Arabic Translation: "عض الرصاصة" (Aad al-risasa)

Issue: The phrase "عض الرصاصة" translates directly to "bite the bullet," which does not have the same idiomatic connotation in Arabic. The intended meaning might be lost, and Arabic speakers might not understand the metaphorical use of "bullet" in this context.

Example 3: "Let the cat out of the bag"

English Idiom: "Let the cat out of the bag"

Intended Meaning: To reveal a secret or disclose something that was meant to be kept confidential.

Literal Arabic Translation: "أطلق القطة من الكيس" (Atlaq al-qitta min al-kees)

Analysis: The literal translation "أطلق القطة من الكيس" means "let the cat out of the bag," which is not an idiomatic expression in Arabic. Arabic speakers may not understand the metaphor, and the translation does not effectively convey the idea of revealing a secret.

Example 4: "Hit the nail on the head"

English Idiom: "Hit the nail on the head"

Intended Meaning: To describe precisely or accurately.

Literal Arabic Translation: "ضرب المسمار على الرأس" (Dharb al-mismaar 'ala al-ra's)

Analysis: The literal translation "ضرب المسمار على الرأس" means "hit the nail on the head," which may not be idiomatic in Arabic. The expression might be unclear, as the metaphor might not have a direct equivalent in Arabic, leading to confusion about its intended meaning.

Example 5: "Spill the beans"

English Idiom: "Spill the beans"

Intended Meaning: To disclose secret information or reveal something that was meant to be kept confidential.

Literal Arabic Translation: "سفك الفاصوليا" (Safka al-fasooliya)

Analysis: The literal translation "سفك الفاصوليا" means "spill the beans," which is nonsensical in Arabic and does not convey the idiomatic meaning of revealing secrets. This highlights a failure to capture the idiomatic nature of the phrase.

٤.2. Inability to Address Word-Order Variations

Here are examples illustrating the issue of word-order variations in machine translation from English to Arabic. These examples highlight how discrepancies in word order between languages can lead to awkward or incorrect translations.

Example 1: "She gives him a book"

English Sentence: "She gives him a book"

Standard English Word Order: Subject-Verb-Object (SVO)

Literal Arabic Translation: "هي تعطيه كتاب" (Hiya ta'tihee kitaab)

Analysis: The literal translation "هي تعطيه كتاب" follows the Subject-Verb-Object (SVO) order as in English. However, in Arabic, while this order is generally acceptable, the omission of the definite article and proper context can make the sentence sound unnatural or unclear. Arabic often prefers more context or a different structure for clarity.

Example 2: "The cat sat on the mat"

English Sentence: "The cat sat on the mat"

Standard English Word Order: Subject-Verb-Object (SVO)

Literal Arabic Translation: "القط جلس على السجادة" (Al-qiṭ jālas 'ala al-sijāda)

Analysis: While the literal translation "القط جلس على السجادة" follows the English SVO order, it may not fully capture the natural phrasing in Arabic. Arabic typically uses a more flexible word order, and in some contexts, a more nuanced structure might be preferred, such as emphasizing the location or the action.

Example 3: "He quickly solved the problem"

English Sentence: "He quickly solved the problem"

Standard English Word Order: Subject-Adverb-Verb-Object (SAVO)

Literal Arabic Translation: "هو بسرعة حل المشكلة" (Huwa bisur'a ḥall al-mushkila)

Analysis: The English adverb "quickly" is translated as "بسرعة" (bisur'a), but its position immediately before the verb may not reflect the natural emphasis in Arabic. In Arabic, adverbs can be placed differently, and the translation might need reordering for more natural flow and emphasis.

Example 4: "The teacher gave the students homework"

English Sentence: "The teacher gave the students homework"

Standard English Word Order: Subject-Verb-Indirect Object-Direct Object (SVIDO)

Literal Arabic Translation: "المعلم أعطى الطلاب الواجب" (Al-mu'allim a'ṭā al-ṭullāb al-wājib)

Analysis: The translation "المعلم أعطى الطلاب الواجب" follows a similar SVIDO order but may not fully capture the emphasis or context in Arabic. Arabic might prefer reordering or additional context for clarity, such as explicitly mentioning "homework" in a way that aligns with common usage in the target language.

Example 5: "She asked him to leave early"

English Sentence: "She asked him to leave early"

Standard English Word Order: Subject-Verb-Object-Infinitive Adverb (SVOIA)

Literal Arabic Translation: "هي طلبت منه أن يغادر مبكرًا" (Hiya ṭalabat minhu an yughadir mubakkiran)

Analysis: In Arabic, while the translation "هي طلبت منه أن يغادر مبكرًا" follows a similar structure, the placement of the adverb "مبكرًا" (mubakkiran) at the end might affect the naturalness of the sentence. Arabic may favor different phrasing to emphasize urgency or context more clearly.

These examples demonstrate how machine translation systems might struggle with word-order variations between English and Arabic, resulting in translations that may be grammatically correct but lack naturalness or clarity.

4.3 Bias in Knowledge Patterns

The reliability issues observed in (AI Translation) can be attributed to biases present in the input data. These biases arise from the training patterns employed by neural networks. Since neural networks learn from the data they are exposed to, any inherent biases in this data can influence the translation process. Consequently, the neural network may produce inaccurate translations due to these biases, resulting in outputs that do not appropriately reflect the intended meaning.

For example, at the outset of the (Republican) Donald Trump's presidency in late 2016, the available data for (AI Translation) would predominantly reference the (Democrat) "US President Barack Obama." Consequently, any occurrence of "US President Trump" would be considered highly unlikely by the system. Given the predominance of data associated with "President Obama," (AI Translation) is trained to default to this former reference when encountering new data, such as "President Trump." This issue arises with any new input not represented in the training data.

When (AI Translation) is tasked with processing information outside its training knowledge, it often yields nonsensical results. This lack of coherence can be particularly evident in translations from English to Arabic, where outputs frequently exhibit inconsistencies, unreliability, and sometimes even absurdities.

Additionally, this highlights the limitations of (AI Translation) in handling various languages. When applied to languages with less representation on the web, such as Arabic, compared to more prevalent languages like English, French, Italian, German, or Spanish the resulting translations often diverge significantly from the intended meaning.

Bias in knowledge patterns can manifest in various ways when AI performs translations compared to human translators. Here are a few examples of how this bias might appear:

a- Cultural Context:

AI Translation:

- **English:** “She is a career woman.”

- **AI Arabic Translation:** "هي امرأة تعمل."

Analysis: The AI might translate this without considering cultural nuances, potentially missing the connotation that being a "career woman" can have positive or neutral connotations in different cultures.

Human Translation:

- **Arabic Translation:** "هي امرأة ناجحة في مسيرتها المهنية."
- **Explanation:** A human translator might add a nuance to reflect the positive connotation of success in a career, which is more culturally sensitive and appropriate.

b- Gender Sensitivity:

AI Translation:

- **English:** "The doctor should follow up with her patient."
- **AI Arabic Translation:** "يجب على الطبيب متابعة مريضه."

Analysis: The AI might default to male terms for professions, even when the gender is unspecified, leading to gender bias.

Human Translation:

- **Arabic Translation:** "يجب على الطبيب/الطبيبة متابعة مريضه/مريضتها."
- **Explanation:** A human translator may make the gender explicit or use a gender-neutral approach if the context requires it.

c- Ambiguity and Nuance:

AI Translation:

- **English:** "She has a lot of experience."
- **AI Arabic Translation:** "لديها الكثير من الخبرة."

Analysis: The AI might not capture the specific context or nuance of the experience being discussed, which could vary significantly depending on the situation.

Human Translation:

- **Arabic Translation:** "تتمتع بخبرة واسعة في هذا المجال."
- **Explanation:** A human translator can tailor the translation to fit the context better, adding specificity that reflects the intended meaning more precisely.

d- Bias in Training Data:

AI Translation:

- **English:** "The scientist presented her research."
- **AI Arabic Translation:** "العالم قدم بحثه."

- **Issue:** If the AI's training data includes a bias toward using male pronouns for professions, it might default to a male term for "scientist," even if the original text specifies female.

Human Translation:

- **Arabic Translation:** "العالمة قدمت بحثها."

- **Explanation:** A human translator will pay attention to gender pronouns and other details, ensuring that the translation accurately reflects the original text's gender specification.

In these examples, AI might struggle with cultural nuances, gender neutrality, idiomatic expressions, ambiguity, and biases presented in its training data. Human translators are generally better at navigating these subtleties and providing contextually appropriate translations.

4.4. Translating in isolation

(AI Translation) intelligent algorithms frequently translate words, phrases, and sentences in isolation, which is a significant limitation. This results in output that is out of context and often unintelligible. Machine translations lack the human ability to understand the context of a given language item before translating it. When (AI Translation) are used to translate a tale, each sentence is translated separately and then combined at the end. This does not make sense.

Translating in isolation can lead to errors when translating from English to Arabic, as context is crucial for understanding the meaning and nuances of a sentence. Here are some examples illustrating how translating phrases in isolation can result in inaccurate or misleading translations:

Example 1: "He's in the bank."

English Sentence: "He's in the bank."

Literal Arabic Translation: "هو في البنك" (Huwa fī al-bank)

Analysis: The literal translation "هو في البنك" (He is in the bank) could be ambiguous. In English, "bank" could refer to a financial institution or the side of a river. Without context, it's unclear whether "bank" refers to a place of financial transactions or a riverbank. If the context was intended to refer to a financial institution, an additional context is needed to confirm the intended meaning.

Example 2: "She can't make it."

English Sentence: "She can't make it."

Literal Arabic Translation: "هي لا تستطيع عملها" (Hiya lā taḥtāṭī‘ ‘amaluhā)

Analysis: The literal translation "هي لا تستطيع عملها" (She can't do it) does not capture the idiomatic meaning of "She can't make it," which generally means "She cannot attend or arrive." The context here is crucial; the phrase could mean she cannot attend an event or complete a task. A more contextually appropriate translation could be "هي لا تستطيع الحضور" (Hiya lā taṣtaṭī‘ al-ḥuḍūr) if referring to attendance, or "هي لا تستطيع الوصول" (Hiya lā taṣtaṭī‘ al-wuṣūl) if referring to arriving.

Example 3: "The meeting was a wash."

English Sentence: "The meeting was a wash."

Literal Arabic Translation: "الاجتماع كان غسلاً" (Al-ijtima‘ kān ghaslan)

Analysis: The literal translation "الاجتماع كان غسلاً" (The meeting was a wash) is confusing because "wash" in this context means that the meeting was a failure or unproductive. Without context, this idiomatic expression does not translate well into Arabic. A more accurate translation that conveys the intended meaning could be "الاجتماع كان فاشلاً" (Al-ijtima‘ kān fāshilan), meaning "The meeting was a failure."

Example 4: "He's on the ball."

English Sentence: "He's on the ball."

Literal Arabic Translation: "هو على الكرة" (Huwa ʿalā al-kurah)

Analysis: The literal translation "هو على الكرة" (He's on the ball) does not convey the idiomatic meaning of being alert or attentive. In Arabic, this phrase would be understood literally, not as an idiom. A more contextually accurate translation could be "هو يقظ" (Huwa yaqẓ), meaning "He is alert."

Example 5: "Give me a hand."

English Sentence: "Give me a hand."

Literal Arabic Translation: "أعطني يداً" (A‘ṭinī yadān)

Analysis: The literal translation "أعطني يداً" (Give me a hand) can be interpreted as asking for a physical hand, rather than the idiomatic meaning of requesting assistance. A more contextually appropriate translation could be "ساعدني" (Sa‘idnī), which means "Help me."

Example 6: "She's a real gem."

English Sentence: "She's a real gem."

Literal Arabic Translation: "هي جوهرة حقيقية" (Hiya jawhara ḥaqīqiyya)

Analysis: The literal translation "هي جوهرة حقيقية" (She is a real gem) might be understood as a literal gem rather than the idiomatic meaning, which implies that she is a valuable or exceptional person. A more suitable translation could be "هي شخص مميز" (Hiya shakhs mumayaz), meaning "She is an exceptional person."

These examples underscore the importance of context in translation. Translating phrases in isolation without considering the broader context can lead to inaccurate or misleading translations.

Additionally, neural networks may fail to recognize a well-publicized international political crisis. Missing up on this opportunity is detrimental to the (AI Translation) reputation and future success.

4.5. Anaphoric errors

(AI Translation) systems' dependence on statistical patterns in training data leads to anaphoric errors in translation.

When anaphora is used as a pronoun, intelligent algorithms may assume gender for distinct occupations. Nurses tend to be female, whereas pilots are often male. This causes the system to use incorrect pronouns even when the context indicates otherwise.

Anaphoric mistakes in translation occur when the AI fails to correctly link pronouns or other referential expressions to their corresponding antecedents, resulting in unclear or incorrect translations. Here are examples illustrating how anaphoric mistakes can occur when translating from English to Arabic:

Example 1: Incorrect Anaphoric Reference

English Sentence: "John said he would call his mother later."

Literal Arabic Translation: "قال جون إنه سيتصل بأمه لاحقاً." (Qāl Jūn innahu sayattaṣil bi-ummihi lāḥiqan.)

Analysis: The pronoun "he" in English refers to "John," and "his mother" refers to John's mother. The literal Arabic translation correctly reflects these references. However, if the AI system were to translate without properly handling anaphora, it might produce a sentence like "قال جون إنه سيتصل بأمه لاحقاً" but with the incorrect understanding of who "his mother" is. For instance, if "his" were incorrectly interpreted to refer to someone else mentioned earlier, the translation might erroneously indicate that "he" (John) is calling someone else's mother.

Correct Translation: "قال جون إنه سيتصل بوالدته لاحقاً." (Qāl Jūn innahu sayattaṣil bi-wālidatihi lāḥiqan.)

This correctly maintains the reference to John’s mother.

Example 2: Misinterpreting Gender

English Sentence: "When Emily arrived, she was happy because it was sunny."

Literal Arabic Translation: "عندما وصلت إميلي، كانت سعيدة لأن الطقس كان مشمساً." (ʿIndamā waṣalat Īmīlī, kānat saʿīda li'anna al-ṭaqs kān mushmisan.)

Analysis: The pronoun "she" refers to Emily, and "it" refers to the weather. If the AI mishandles the anaphoric reference, it might incorrectly link "it" to "Emily," leading to a translation that implies Emily was sunny rather than the weather.

Correct Translation: "عندما وصلت إميلي، كانت سعيدة لأن الطقس كان مشمساً." (ʿIndamā waṣalat Īmīlī, kānat saʿīda li'anna al-ṭaqs kān mushmisan.)

Here, the pronoun "it" is correctly interpreted as referring to the weather.

Example 3: Incorrect Reference to Multiple Antecedents

English Sentence: "Alice told Sarah that she would help her with the project."

Literal Arabic Translation: "أخبرت أليس سارة أنها ستساعدها في المشروع." (Akhtarat Ālīs Sārā annahā sataʿāwinuhā fī al-mashrūʿ.)

Analysis: In this sentence, "she" refers to Alice and "her" refers to Sarah. If the AI does not properly handle the anaphoric references, it might misinterpret "she" as referring to Sarah, resulting in " أخبرت سارة أليس أنها ستساعدها" (which would be incorrect, suggesting Sarah is helping Alice).

Correct Translation: "أخبرت أليس سارة أنها ستساعدها في المشروع." (Akhtarat Ālīs Sārā annahā sataʿāwinuhā fī al-mashrūʿ.)

This correctly identifies that Alice is the one who will help Sarah with the project.

Example 4: Ambiguous Pronoun Resolution

English Sentence: "Tom and Jerry were at the park. He found it amusing."

Literal Arabic Translation: "توم وجيري كانا في الحديقة. لقد وجده مسلياً." (Tūm wa Jīrī kānā fī al-ḥadīqa. Laqad wajadahū musliyyan.)

Analysis: In this case, "he" refers to either Tom or Jerry, and "it" refers to the park. If the AI does not correctly resolve the anaphoric references, the translation might ambiguously suggest that "he" (Tom or Jerry) found "it" (the park) amusing, without clear identification.

Correct Translation: "توم وجيري كانا في الحديقة. وجد توم ذلك مسلياً." (Tūm wa Jīrī kānā fī al-ḥadīqa. Wajad Tūm dhālika musliyyan.)

Here, specifying "Tom" clarifies the reference, ensuring that Tom found the park amusing.

Example 5: Confusing Pronoun References

English Sentence: "The manager asked the assistant if he could handle the task."

Literal Arabic Translation: "سأل المدير المساعد إذا كان يستطيع التعامل مع المهمة." (Sa'ala al-mudīr al-musaʻid ʿidhā kān yastaṭīʿ al-taʻāmul maʻa al-mihma.)

Analysis: The pronoun "he" refers to the assistant, but if the AI translates this without correctly handling anaphora, it might incorrectly interpret "he" as referring to the manager or someone else mentioned previously.

Correct Translation: "سأل المدير المساعد إذا كان يستطيع التعامل مع المهمة." (Sa'ala al-mudīr al-musaʻid ʿidhā kān yastaṭīʿ al-taʻāmul maʻa al-mihma.)

This ensures that the reference to "he" is correctly understood as the assistant.

These examples underscore the importance of correctly handling anaphora in AI translation to avoid misinterpretation and ensure that the translated text accurately conveys the intended meaning.

5- Results:

a- Inherent Limitations of AI Translations

(AI Translation) systems are typically designed to process one sentence at a time, which limits their ability to handle longer documents or broader contexts. This constraint hinders the coherence of machine-generated translations and prevents them from achieving a quality comparable to human translation. Additionally, encoding a comprehensive understanding of the world into neural networks is challenging, leading to translations that often lack common sense and contextual awareness. This limitation is evident in the various examples discussed throughout this paper.

b- Absence of a Comprehensive Self-Evaluation Mechanism

Currently, (AI Translation) systems lack an integrated mechanism for self-evaluating translation quality. Consequently, assessing the accuracy of these translations requires human intervention. The BLEU (Bilingual Evaluation Understudy) score is a commonly used metric for evaluating the quality of (AI) translations (Papineni et al., 2002). BLEU compares (AI Translation) outputs with human translations of similar content. However, obtaining comparable human translations for most language pairs is often impractical.

Various evaluation techniques in machine translation (Reeder, 2001) aim to encompass different facets of human assessment, such as adequacy, fidelity, and fluency, as detailed by Hovy (1999) and White & O'Connell (1994).

Following the adoption of neural translation models by Google, CEO Sundar Pichai reported a significant improvement in translation accuracy, with a BLEU score increase from 3.694 to 4.263, and noted that "human quality was only a step away at 4.636" (as quoted in Pan, 2016). However, the BLEU score is not infallible and is subject to limitations inherent in its comparison-based methodology, which is neither fully independent nor free from anomalies.

c- Evaluation of BLEU Score and Human Translation

The BLEU (Bilingual Evaluation Understudy) score assesses machine translation (MT) outputs by comparing them to human-translated texts. A higher BLEU score indicates greater similarity between the machine-generated output and human translations. However, this metric is constrained by its reliance on exact matches and the necessity of human-generated reference translations. Consequently, it is inaccurate to describe BLEU's evaluation as truly human-like or close to human-like.

d- (AI Translation): Progress and Challenges

We evaluated the current neural machine translation (AI Translation) systems were evaluated by Google across various types of content, including social idioms and contemporary political issues, to gauge the current state of neural translation technology. The results, detailed in the preceding sections, reinforce the notion that, (AI Translation) has significant limitations and is far from being a viable substitute for human translation.

The examples of English-to-Arabic translations provided by AI Translation tools reveal several issues, ranging from outright inaccuracies to highly unreliable or incomprehensible results, and in some cases, translations that are downright embarrassing. These issues highlight the (AI translation) systems' frequent errors, including omissions of semantic content, misinterpretation of phrases, and syntactic alterations, leading to outputs that are often unintelligible.

Given these challenges, it is crucial to rigorously assess the quality and accuracy of (AI Translations) translations. The BLEU algorithm used for this evaluation is not free from anomalies and limitations, as previously discussed.

Despite these challenges, the field of (AI Translation) is advancing rapidly. Researchers and practitioners at leading institutions and companies are continuously working to address fundamental issues related to memory, data bias, common sense knowledge, and the accuracy of outputs. This ongoing development promises to mitigate many of the current limitations and improve the overall reliability of (AI Translation) systems.

e- Advances in AI Translation

Significant progress has been made in enhancing the effectiveness and efficiency of (AI Translation). Notable advancements include Google's Transformer and Salesforce's Quasi-Recurrent Neural Networks, which utilize various data processing techniques ranging from sequential to parallel processing. These innovations have led to more efficient data training and, consequently, more accurate translations.

Additionally, groundbreaking research continues to emerge in this field. For instance, Harvard's OpenNMT, an open-source neural machine translation framework developed in LuaTorch, PyTorch, and TensorFlow, aims to integrate advanced data processing methods. Similarly, DeepL, created by a former Google scientist, reportedly offers substantial improvements over Google Translate's outputs.

Moreover, new multilingual enterprise solutions are continually being developed, such as Microsoft Translator, Facebook Translate, and Baidu APK. While advancements in deep learning are expected to drive (AI Translation) evolution, it remains unlikely to fully replace human translation in the near future. Despite increasing access to data and ongoing improvements in neural network algorithms, challenges such as reliability in translating more than single sentences persist. (AI Translation) may effectively provide quick summaries, but for contexts requiring precise language, such as literature, diplomacy, and formal communication, it still falls short of reliability. Additionally, (AI Translation) struggles with languages and language combinations that are less common in the digital domain.

However, there have been notable achievements in specific language pairs. For example, Microsoft claims to have achieved human-level accuracy and quality in translating between Chinese and English. Researchers at Microsoft utilized Dual Learning and Deliberation Networks to train their model on approximately 2,000 sentences from the newstest2017 dataset (Dar, 2018). Two innovative techniques, Joint Training and Agreement Regularization, were employed to further enhance translation quality. Joint Training involved iterative translation between Chinese and English, while Agreement Regularization ensured that the system could translate effectively in both directions (left-to-right and right-to-left). Similar results from both directions were used to assess translation reliability. External bilingual experts also evaluated the machine translations by comparing them with human translations. Although these developments are promising, they have yet to be tested with real-time news stories.

f- AI Translation and its Challenges

(AI Translation) has demonstrated notable success, particularly with language sets that involve complex morphology and significant word reordering, as highlighted by Luong and Manning (2015) and Bentivogli et al. (2016). (AI Translation) employs an end-to-end encoder-decoder framework to process a sentence or language input and generate its translation. However, this traditional approach, as noted by Tu et al. (2017), presents two significant issues: "Translations generated by (AI Translation) systems often lack adequacy" and "The likelihood objective is suboptimal in decoding."

The issue of adequacy arises when AI translations frequently exhibit repetition of target words for certain segments of a sentence while neglecting other parts, resulting in over-translation and under-translation (Tu et al. 2016). The problem of "suboptimal likelihood objective in decoding" relates to the beam search algorithm used by (AI Translation) to maximize likelihood. This feature often falls short in large decoding spaces because it captures only unidirectional dependencies from the source to the target language, which does not align well with translation adequacy (Li & Jurafsky 2016; Shen et al. 2016).

To address these challenges, Tu et al. (2017) propose a new unified framework known as the encoder-decoder reconstructor. This framework aims to expand the decoding space, potentially enhancing the quality and accuracy of (AI) translations.

6- Conclusion

The notion of machines replacing human intelligence is indeed compelling; however, the complete replacement of human bilingual expertise with current neural networks and deep learning technologies seems unlikely in the near future. Instead, it is expected that translation tasks will increasingly be shared between machines and humans, depending on the specific requirements of the task.

The pervasive presence of speech and language-processing technologies in daily life makes it challenging to avoid their influence in translation and other language-related activities. For instance, Jurafsky et al. (2009) highlight how "Google provides cross-language information retrieval and translation services, allowing users to submit queries in their native language and receive translated results in another language. Google translates the query, retrieves the most relevant pages, and then translates them back into the user's native language." This functionality is made possible by advanced (AI Translation) systems utilized by various media platforms for real-time translations.

Despite these successes, the limitations of (AI Translation) compared to human translation are evident. As Douglas Hofstadter (2018) aptly notes, "When, one day, a translation engine crafts an artistic novel in verse in English, using precise rhyming iambic tetrameter rich in wit, pathos, and sonic verve, then I'll know it's time for me to tip my hat and bow out."

References

Bahdanau, D., Cho, K., & Bengio, Y. (2015). Neural machine translation by jointly learning to align and translate. *International Conference on Learning Representations*.

Bentivogli, L., Bisazza, A., Cettolo, M., & Federico, M. (2016). Neural versus phrase-based machine translation quality: a case study. *Empirical Methods in Natural Language Processing (EMNLP)*.

Berger, A., Della Pietra, S. A., & Della Pietra, V. J. (1996). A maximum entropy approach to natural language processing. *Computational Linguistics, 22*(1), 39–71.

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research, 3*(5), 993–1022.

Cho, K., van Merrienboer, B., Gülçehre, Ç., Bougares, F., Schwenk, H., & Bengio, Y. (2014). Learning phrase representations using RNN encoder-decoder for statistical machine translation. *Conference on Empirical Methods in Natural Language Processing*.

Dar, P. (2018). Microsoft's language translation: AI has reached human levels of accuracy. *Analytics Vidhya*. Available at: https://www.analyticsvidhya.com/blog/2018/03/microsofts-claimslanguage-translation-ai-reached-human-levels-accuracy/

Hajic, J. (1998). Building a Syntactically Annotated Corpus: The Prague Dependency Treebank (pp. 106–132). Karolinum.

Hao, K. (2018). Baidu's Chinese-to-English translator finishes your sentence for you. *MIT Technology Review*. Available at: https://www.technologyreview.com/the-download/612338/baidus-chineseto-english-translator-finishes-your-sentence-for-you/

Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*.

Hofstadter, D. (2018). The Shallowness of Google Translate. *The Atlantic*. Available at: https://www.theatlantic.com/technology/archive/2018/01/theshallowness-of-google-translate/551570/

Hong, C. (2018). Construction of corpus in Artificial Intelligence age. *MATEC Web of Conferences, 175*, 03037. Available at: https://www.matecconferences.org/articles/matecconf/abs/2018/34/matecconf_ifcaeiot2018_03037/matecconf_ifcae-iot2018_03037.html

Hovy, E. H. (1999). Toward finely differentiated evaluation metrics for machine translation. *Proceedings of the Eagles Workshop on Standards and Evaluation*, Pisa, Italy.

Hunt, G. H., Bennett, M. A., Carboni, J. N., & Pantanelli, G. P. (2017). Method to resolve the meaning of a body of natural language text using artificial intelligence analysis in combination with semantic and contextual analysis. *United States Patent Application Publication*, US20170286398A1. Available at: https://patentimages.storage.googleapis.com/f6/50/ee/e16be81bf43296/US20170286398A1.pdf

Johnson, M., Schuster, M., Le, Q., Krikun, M., Wu, Y., Chen, Z., Thorat, N., Viégas, F., Wattenberg, M., Corrado, G., Hughes, M., & Dean, J. (2017). Google's multilingual neural machine translation system: enabling zero-shot translation. Available at: https://arxiv.org/pdf/1611.04558.pdf

Jurafsky, D., & Martin, J. (2009). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. New Jersey: Prentice-Hall.

Kilgarriff, A., & Palmer, M. (Eds.). (2000). Computing and the Humanities: Special Issue on SENSEVAL, Vol. 34. Kluwer.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*, 436-444. Available at: https://www.nature.com/articles/nature14539

Li, J., & Jurafsky, D. (2016). Mutual information and diverse decoding improve neural machine translation. *NAACL*.

Luong, M. T., & Manning, C. D. (2015). Stanford neural machine translation systems for spoken language domains. *IWSLT*.

Luong, M. T., Pham, H., & Manning, C. D. (2015). Effective approaches to attention-based neural machine translation. *Empirical Methods in Natural Language Processing (EMNLP)*.

Marcus, M. P., Santorini, B., & Marcinkiewicz, M. A. (1993). Building a large annotated corpus of English: The Penn Treebank. *Computational Linguistics, 19*(2), 313–330.

Marr, B. (2018). Will machine learning AI make human translators an endangered species? *Forbes*. Available at: https://www.forbes.com/sites/bernardmarr/2018/08/24/will-machinelearning-ai-make-human-translators-an-endangeredspecies/#6b63ae483902

Massey, G., & Ehrensberger-Dow, M. (2017). Learning machines and teaching humans. *International Conference: Points of View in Translation and Interpreting*, Kraków, Poland, June 22–23. Publikationen Angewandte Linguistik. Available at: https://digitalcollection.zhaw.ch/handle/11475/3278

Palmer, M., Kingsbury, P., & Gildea, D. (2005). The Proposition Bank: an annotated corpus of semantic roles. *Computational Linguistics, 31*(1), 71–106.

Palmer, M., Fellbaum, C., Cotton, S., Delfs, L., & Dang, H. T. (2001). English tasks: all-words and verb lexical sample. *Proceedings of Senseval-2: 2nd International Workshop on Evaluating Word Sense Disambiguation Systems*, Toulouse, France, pp. 21–24.

Pan, H. M. (2016). How BLEU measures translation and why it matters. *Slator: Language Industry Intelligence*. Available at: https://slator.com/technology/how-bleu-measures-translation-and-why-itmatters/

Papineni, K., Roukos, S., Ward, T., & Zhu, W. (2002). BLEU: a method for automatic evaluation of machine translation. *Proceedings of the 40th Annual Meeting of the Association for Computational Linguistics (ACL)*, Philadelphia, July 2002, pp. 311–318. Available at: https://www.aclweb.org/anthology/P02-1040.pdf

Pring-Mill, D. (2018). Why hasn't AI mastered language translation. *SingularityHub*. Available at: https://singularityhub.com/2018/03/04/whyhasnt-ai-mastered-languagetranslation/#sm.000000eb79kzgld29rvnyvdl75493

Pustejovsky, J., Hanks, P., Saurí, R., See, A., Gaizauskas, R., Setzer, A., Radev, D., Sundheim, B., Day, D. S., Ferro, L., & Lazo, M. (2003). The TIMEBANK corpus. *Proceedings of Corpus Linguistics 2003 Conference*, pp. 647–656. UCREL Technical Paper number 16.

Reeder, F. (2001). Additional mt-eval references. *Technical report, International Standards for Language Engineering, Evaluation Working Group*. Available at: http://isscowww.unige.ch/projects/isle/taxonomy2/

Shen, S., Cheng, Y., He, Z., He, W., Wu, H., Sun, M., & Liu, Y. (2016). Minimum risk training for neural machine translation. *Association for Computational Linguistics (ACL)*.

Sutskever, I., Vinyals, O., & Le, Q. V. (2014). Sequence to sequence learning with neural networks. *Advances in Neural Information Processing Systems*, pp. 3104–3112.

Tu, Z., Liu, Y., Shang, L., Liu, X., & Li, H. (2017). Neural machine translation with reconstruction. *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence*, pp. 3097–3103. Available at: AAAI-17

Tu, Z., Lu, Z., Liu, Y., Liu, X., & Li, H. (2016). Modeling coverage for neural machine translation. *Association for Computational Linguistics (ACL)*.

White, J. S., & O'Connell, T. (1994). The ARPA MT evaluation methodologies: evolution, lessons, and future approaches. *Proceedings of the First Conference of the Association for Machine Translation in the Americas*, pp. 193–205, Columbia, Maryland.

Wu, Y., Schuster, M., Chen, Z., Le, Q. V., Norouzi, M., Macherey, W., Krikun, M., Cao, Y., Gao, Q., Macherey, K., Klingner, J., Shah, A., Johnson, M., Liu, X., Kaiser, Ł., Gouws, S., Kato, Y., Kudo, T., Kazawa, H., Stevens, K., Kurian, G., Patil, N., Wang, W., Young, C., Smith, J., Riesa, J. J., Rudnick, A., Vinyals, O., Corrado, G., Hughes, M., & Dean, J. (2016). Google's neural machine translation system: bridging the gap between human and machine translation. Available at: arXiv preprint arXiv:1609.08144.

Zanaty, Dina Gaber. (2024). When Translating from Arabic to English and vice versa, is Google Bard a Trustworthy Tool?. مجلة وادي النيل للدراسات والبحوث الإنسانية والاجتماعية والتربويه, 44(44), 205-236. doi.org/10.21608/jwadi.2024.387211

Zhou, S. (2018). Has AI surpassed humans at translation? Not even close! *Skynet Today*. Available at: https://www.skynettoday.com/editorials/state_of_nmt

Zong, Z. (2018). Research of the relations between machine translation and human translation. *Journal of Physics: Conference Series, 1087*. IOP Publishing. Available at: http://iopscience.iop.org/article/10.1088/1742-6596/1087/6/062046/pdf