Evolving Fuzzy Models for Automated Translation

Alina Ţenescu[1], Radu-Emil Precup[2], Nicuşor Minculete[3]

[1]Faculty of Letters, University of Craiova, Str. A. I. Cuza 13, 200585 Craiova, Romania; E-mail: tenescu.alina@ucv.ro

[3]Department of Automation and Applied Informatics, Politehnica University of Timisoara, Bd. V. Parvan 2, 300223 Timisoara, Romania
E-mail: radu.precup@upt.ro

[3]Faculty of Mathematics and Informatics, Transilvania University of Brasov, Str. Iuliu Maniu 50, 500091 Brasov, Romania; E-mail: minculete.nicusor@unitbv.ro

Abstract: This paper targets two goals. First, it analyzes the most common errors in automated translation from French to English and from English to French performed by a statistical and a hybrid translation engine with the help of the evaluation metric SAE J2450. The test of concordance is applied in order to study the agreement between the original text and its retroversion within the same translation memory software. Seven categories of primary errors are considered, which cover the fields of terminology, semantics, structure, orthography, punctuation and completeness. Second, evolving fuzzy models are developed, which give the overall paragraph score using the seven categories of primary errors as inputs. The fuzzy models permit the users to establish the accuracy of translation and to grade the quality of translations resulted from the reintroduction of the result of translation in the same software application. They also allow the comparison of two popular translation memory programs, namely Google Translate (GT) and Systran, in the framework of the issues of concordance in translation and artificial learning.

Keywords: accuracy; automated translation; concordance; errors; evolving fuzzy systems

1   Introduction

Nowadays, we are witness to an unquestionable reality: the extent and dimensions of informative and formative space are increasingly influenced by the Internet, by computerization and automation which pervade the users' cognitive sphere, directing the contents and orienting the global organization of their ideas. The Internet and computer systems are becoming determining factors of the new dimensions of educational area – revealing the intentions, the content, the means

and the instruments of education – and provide access to a wide range of contents, information and current trainings, which prove themselves necessary all the more so as we all participate to the multidimensional process of globalization of education. In order to implement a new teaching program (for example, in foreign languages or translation) which makes use of traditional methods, but also of the original means made available by new technologies, the teaching staff (teaching to students in translation and interpreting) need training and specialization in three domains: specialization in the domain of pedagogy and education, specialization in the scientific domain linked to the informational content of teaching and specialization in the domains of informatics and programming. We also have to take into account the fact that the new generation of students, "the digital natives", answers more easily and effectively to a novel way of learning, combining the classic classroom sessions with independent learning and with computer-based training (CBT) and web-based training (WBT). The education program based on blended learning helps students in foreign languages and translation to develop digital skills and increases the level of knowledge and handling of online tools of translation engines, in parallel with the improvement of their textual competence and the enhancement of information and technology literacy.

The first goal of this paper is to analyze the most frequent errors in automatic translation. These errors can be identified by students in career training (translation and interpreting) and by professional translators as they carry out translation of a text from French to English ("theme") and from English to French ("version"), following the automatic processing of a paragraph of academic text by two translation memory software applications, namely Google Translate (GT) and Systran. We will then apply the concordance test ("goodness of fit" test, the "adjustment" test) on texts which are translated from French to English and from English to French by each translation software mentioned above, in order to establish the degree of concordance between the original text T1, which is translated, from English to French ("version") and the text T3, which represents the result obtained after the translation into French of T1 is reintroduced within the same translation memory software and retranslated into English.

In the context of the first goal, the text T1 in English will be translated into French with the help of an online translation memory software application (for example, GT or Systran). We obtain the text T2 in French. In a second phase, we introduce the text T2 in the same application and we translate it into English. We note the result with T3 and we compare T1 to T3, establishing the degree of concordance between them with the aid of a statistical test. We will use a statistical test to study the agreement or the concordance of the translation to the original text and to compare T1 with T3.

The analysis of the most frequent errors in automatic translation leads to seven categories of primary errors, i.e. those which cover the fields of terminology, semantics, structure, orthography, punctuation and completeness. The second goal of this paper is to develop fuzzy models, which give the Overall Paragraph Score

(OPS) using these seven categories of primary errors as inputs. The evolving fuzzy models are built upon the results obtained by the Process Control group of the Politehnica University of Timisoara, Romania, and reported in [1-6]. Four evolving fuzzy models are developed for the translation from English to French of the data obtained using GT and Systran in terms of employing either the Recursive Least Squares (RLS) algorithm or the weighted Recursive Least Squares (wRLS) algorithm to update the rule consequents of the fuzzy models. The basic version of the incremental online identification algorithm is implemented according to [6] using the software support specific to eFS Lab [7, 8]. Evolving fuzzy models for the translation from French to English are not developed as the translation is accurate enough to produce low errors that create numerical problems in the identification algorithm.

The results of the new ideas expressed as the two goals are important as our final goal is to propose a course or a path making use of an original method and of a pedagogical software tool that makes possible for students in translation and professional translators to extend and detail research on the existing differences between the degrees of accuracy and quality of translations furnished by different language translation software applications, as well as on the recurrent errors with each translation software. Since translated texts or paragraphs are far from being correct or perfect, the process of automated translation as imperfect model obliges students not only to discriminate the wrong elements found in translation, but also to critically analyze and improve their translator skills by constantly having to refine, revise and proofread an instant automated translation. The use of nonlinear models as fuzzy ones can be beneficial with this regard.

This paper is organized as follows: the analysis of translation results with focus on GT and Systran is carried out in the next section. The metamorphosis of a text through translation (quality and evaluation) is first treated, and a discussion on the choice of automatic translation engines and evaluation metric for translation quality is included. Section 3 is dedicated to the development of the new evolving fuzzy models. The results are validated on real data. The conclusions are highlighted in Section 4.

2   Analysis of Translation Results

Nowadays, a translator is not only he who deciphers the message issued by a transmitter, but also he who recodes it and renders it comprehensible to the receiver. This is though not possible without touching and altering a little bit the content of the message, sometimes at the expense of its quality, value and concordance by comparison to the source. As shown in [9], Kumarajiva asserted that translation resembles, in fact, "to already masticated food that will be given to him who cannot chew it himself". Automated translation would then be

tantamount to victuals chewed by the mouth of an artificial organism whose aim would be to guide the human operator in his work as "validator", as verifier and checker of linguistic solutions generated by the implementation of mathematical algorithms, or, according to Toudic et al. [10], as "post-editor" or as proof-reader who edits, revises, corrects and removes errors from translations made by an automated engine, an engine with translation memory, so as to achieve a certain degree of acceptability and quality. Nevertheless, we can easily become aware of the fact that such food does not have the original taste and aroma anymore.

Bowker and Ehgoetz [11] insist on the fact that automated or machine translation (MT) entails the shift from the concept of "absolute quality" to that of threshold of acceptability of translation by the final beneficiary. It is obvious that this questions the notion of relativity of quality, defined by relationship to the level of acceptability of the product for this final recipient.

Today, the emergence of new media and the development of automated translation and of translation memory tools have changed the perspective on the notion of quality in translation. Green et al. argue that this is not anymore focused on how faithful or unfaithful the translation is by relationship to the source-text [12]. On the contrary, it is re-centered to take into account the final usage or the particular use of the translated text.

While, as stated in [10], the tools of computer-assisted translation and translation memories diminish the final responsibility of the translator as far as the quality of the resulting translation is regarded, the quality of the translated text is established by the quality of translation machines, of pre-translated segments, of terminological databases, as well as by the quality of the revision work lead by tools of evaluation of translation quality and by the human post-editor, contends Koehn [13]. Bar-Hillel draws attention on this partnership between the translation automaton and the post-editor, partnership whose major problem is, according to him, "the region of optimality in the continuum of possible divisions of labor" [14].

Though the evaluation of automated translation systems and the evaluation of translation quality prove to be objects of infinite debates and difficult issues to solve, our objective is to evaluate the quality and the concordance (with the source) of the translation produced by systems using fully automatic translation, which does not involve any human participation or revision. Aside from these automated systems, there are numerous tools and instruments supporting translation, such as electronic dictionaries, whose evaluation is achieved, according to Ryan [15], with the participation of individuals, and which relies upon the evaluation measures of the human-machine interface and on the techniques of evaluation of usability, as argue Hartley and Popescu-Belis [16].

This section is focused on the manner to assess, measure and compare the quality of translations provided by two automatic translation systems (GT and Systran), translations (from English to French and from French to English) deemed

completed and used as such by different users, or by other machine translation systems or software applications. We will first identify the most recurrent errors in the translation from English to French and from French to English of a paragraph of academic text processed by two professional machine translation systems (GT and Systran), using the translation quality metric SAE J2450. In a second stage, we will compare, via the test of concordance ("goodness of fit" test), the quality of translations resulted from the reintroduction of the result of translation in the same software application as well as their concordance with the source-text.

Online translators are based on several approaches of automated translation. These approaches are grouped in four main categories: 1) rule-based machine translation (RBMT) which operates in a more complex way than word-to-word translation and applies linguistic rules in three stages (analysis, transfer, generation), allowing different words to be placed in different places, according to the context, as demonstrate Nirenburg [17] and Costa-Jussa et al. [18]; 2) statistical machine translation (SMT) which is a complex form of word translation, word-based or, more recently, phrase-based, using statistical weights in order to choose the most likely translation of a given word. SMT uses probability and adopts a learning algorithm of the translation produced by humans, show Brown et al. [19]; 3) example and analogy-based machine translation which makes use of bilingual corpora with parallel texts as its main knowledge source and is thus based on previously seen examples in these parallel corpora, according to Nagao [20] and Turcato et al. [21]; 4) hybrid machine translation which blends fundamental elements of the rule-based machine translation system with those of the statistical-based system and whose result depends on the quality of the alignment of candidate translations to the source-sentence and will bring the literal sense over into the statistical result, show Boretz [22] and Hunsicker et al. [23], thus taking advantage of the fluidity of statistical machine translation and of the accuracy of rule-based machine translation, asserts Drexler [24]. Hybrid translation is performed in two stages: first, the linguistic analysis of the text in the source language, contends Rubino [25]; secondly, we pass to the target-language, through non-linguistic approaches, by the act of translating sub-sentential segments, while lexical selection in the target-language is achieved, in the second phase, through a language model.

As each approach of automated translation has a different principle and mode algorithm, the results of translation will be different. Given the opposed character of approaches to automated translation, we have decided to undertake a comparative study of translations provided through the automated translation by a statistical machine translation software (GT) and by a hybrid automated translation software (Systran). We have chosen the statistical automated translator GT since it is one of the most popular online translation software applications, and all the more so as it continuously increases the list of language pairs for which it establishes parameters of translation memory, as well as for its easy access, usability and applicability. The second automated translator, Systran, is built on a

hybrid translation software application, one of the most well-known using hybrid machine translation. The reason behind its choice lies in the old kinship with GT, namely GT initially used a rule-based machine translation of Systran, since Systran was one of the first companies to develop and use RBMT systems and is, nowadays, one of the most well-known to use hybrid MT.

Green et al. [26] and Madsen [27] have argued that even though with automatic translation it is impossible to achieve precise translation and though automated translators commit errors, their relevance and utility in our lives are undeniable. Therefore, we should focus on how these automated translators face issues raised by different language pairs, language grammars and certain grammatical structures, on how they handle and tackle differences in linguistic typologies, translation of idioms and colloquialisms, and on how they draw out responses to language change across time and to linguistic, cultural and social barriers and uncertainties.

If one of the objectives of this section is to comprehend the level of performance of the two approaches (SMT and HMT) in automated translation of academic texts, we also want to reveal the weaknesses and the strengths that can affect the performance of the two online translation software applications. A second objective would be to study and analyze the typical errors in the translation of academic texts from English to French and from French to English, using SAE J2450 translation quality standard, in order to discover which of the two approaches is more accurate and precise. Using the categories established by the SAE J2450 evaluation metric for quality of translation, we manually evaluated the results of translations provided by the two software applications and we calculated the score, expressed as OPS, indicating the performance of HMT and SMT. The score represents a specific case of translation.

If we compare the results of translations from French to English and from English to French, as well as the results of translations resulted from the reintroduction of the result of translation in the same software application, we highlight the advantages and disadvantages of the two approaches, taking into account indicators such as the accuracy of translation at several levels, the performance, the flexibility and the adaptation to changes. We also show the relationship between the conclusions of the manual evaluation of results of automated translation with each of the two software applications and the technical procedures of each of the two approaches.

Although, according to Papineni et al [28], the evaluation of automated translation is sometimes considered a sort of "black art" or occult science, we have nevertheless chosen the SAE J2450 translation quality metric because it has the advantage of being easy to use and because it can be applied irrespective of the source-language or target-language. Another advantage is represented by its capacity to try and solve the issue of ambiguities by way of two meta-rules: 1) "When an error is ambiguous, always choose the earliest primary category",

2) "When in doubt, always choose 'serious' over 'minor'" (SAE J2450, [29]), choice which would ensure, according to Secară [30] a greater coherence and "consistency of classification of errors across evaluators".

The evaluation metric comprises four sub-divisions: 1) seven categories of primary errors which cover the fields of terminology, semantics, structure, orthography, punctuation, completeness; 2) two secondary sub-categories: serious error, minor error; 3) two meta-rules which help the evaluators in their decision, in the case of ambiguity; 4) numeric weights for each primary category and sub-category. For the first sub-division, there are recognized seven categories of errors taken into account in the evaluation of translation quality. The overall aim is to generalize, which renders a scheme of general categories of errors with SAE J2450. The errors are first weighed and afterwards classified as 'serious' or 'minor' (SAE, 2001) [29]. The weights of the seven categories of primary errors categorized as serious errors/minor errors are [29]: Wrong Term (WT): 5/2, Syntactic Error (SE): 4/2, OMission (OM): 4/2, Word Structure and Agreement error (SA): 4/2, MisSPelling (SP): 3/1, Punctuation Error (PE): 2/1, and Miscellaneous Error (ME): 3/1.

In this section, in order to obtain the final note for the quality of translation in the target-language, we calculated, in the same manner as the one established by SAE standard, the sum of numerical values of the totality of committed errors and then we divided it by the total number of words in the source-text. Hartley and Popescu-Belis [16] assert that we have thus the "occasion to calibrate in order to favor" either the correct grammatical form or the terminological correctness.

The analysis allowed us to identify the typical errors in the translation from English to French and from French to English performed by the two translation memory programs GT and Systran. We tested not only the performance of each software application, but also the precision of translations provided by these automated translation applications. We chose as corpus of study two paragraphs of recent academic texts:

T1 [31]: "Information literacy has been a tremendous "win" for academic libraries. But it risks becoming, looking back, also a symbol of a great loss. If we do not refocus our efforts on the educational, cultural, and technological shifts in which "information literacy" per se becomes a somewhat arbitrary label for the very stuff of learning and information discovery in today's academic (and larger) world, we will have won the battle but lost the campaign".

T4 [32]: "Une caractéristique de ces différentes approches de l'utilisabilité consiste à étendre le champ d'application de l'ergonomie aux produits et aux technologies interactives de la vie courante. Non seulement parce que ceci constitue un nouveau champ de débouchés, mais surtout parce qu'avec le renouvellement rapide des produits, il existe une relation étroite entre les manières de les consommer et les manières de les produire. À l'expression "comprendre le travail pour le transformer", on pourrait rajouter "comprendre les modes de

relations aux produits pour transformer le travail" et "comprendre les relations aux produits pour agir sur l'expérience".

As far as the paragraphs are concerned, the one in English is an excerpt from an article on information literacy [31], while the one in French is an excerpt from a study on the acceptability of new technologies [32]. We focused on the comparison between the source-paragraph and the translated result. In the choice of paragraphs selected from academic texts, we used as criterion of selection their provenance from a credible source and their relatively recent publication (during the last six years). Some results of the translation are given as follows:

T2 (translated result using GT): maîtrise de l'information a été un "gagnant" énorme pour les bibliothèques universitaires. Mais il risque de devenir, en regardant en arrière, aussi un symbole d'une grande perte. Si nous ne recentrer nos efforts sur les changements éducatifs, culturels et technologiques dans lesquels "information literacy" en soi devient une étiquette quelque peu arbitraire pour la substance même de l'apprentissage et de découverte de l'information dans le monde académique (et plus) d'aujourd'hui, nous aurons gagné la bataille, mais a perdu la campagne.

T3 (reintroduction of the translated result in the same software application, GT): information literacy has been a "win" huge for academic libraries. But it may become, looking back, as a symbol of a great loss. If we focus our efforts on educational change, in which cultural and technological "information literacy" itself becomes a somewhat arbitrary label for the substance of learning and discovery of information in the academic world (and more) today, we have won the battle but lost the campaign.

T2' (translated result using Systran): L'instruction de l'information a été une "victoire" énorme pour les bibliothèques scolaires. Mais elle risque devenir, regardant en arrière, aussi un symbole d'une grande perte. Si nous ne refocalisons pas nos efforts sur les décalages éducatifs, culturels, et technologiques dans lesquels la "instruction de l'information" devient en soi un label quelque peu arbitraire pour la substance même de l'étude et de la découverte de l'information en monde scolaire (et plus grand) d'aujourd'hui, nous aurons gagné la bataille mais aurons perdu la campagne.

T3' (reintroduction of the translated result in the same software application, Systran): The instruction of information was an enourmous "victory" for the school licraries. But it risks to become, looking at behind, also a symbol of a great loss. If we do not refocalisons our efforts on the shifts educational, cultural, and technological in which "instruction of information" becomes in oneself a somewhat arbitrary label for the substance even of the study and of the discovery of information in world school (and larger) of today, we will have gained the battle but will have lost the countryside.

T5 (translated result using GT): A characteristic of these different approaches to usability is to extend the scope of ergonomics products and interactive technologies in everyday life. Not only because this is a new field of opportunities, but mainly because with the rapid turnover of products, there is a close relationship between ways of consuming and ways of producing them. In the term "understand the work to transform the" one might add "understand the products to relationship patterns to transform the work" and understand the relationships to products to act on experience.

T6 (reintroduction of the translated result in the same software application, GT): Une caractéristique de ces différentes approches de la facilité d'utilisation est d'étendre le champ d'application des produits d'ergonomie et technologies interactives dans la vie quotidienne. Non seulement parce que cela est un nouveau champ de possibilités, mais surtout parce que le renouvellement rapide des produits, il existe une relation étroite entre les modes de consommation et les moyens de les produire. Dans le terme "comprendre le travail de transformer" le on pourrait ajouter "comprendre les produits à des modèles relationnels pour transformer le travail" et comprendre les relations aux produits d'agir sur l'expérience.

T5' (translated result using Systran): A characteristic of these various approaches of the utilisability consists in extending the field of application of ergonomics to the products and interactive technologies of the everyday life. Not only because this constitutes a new field of outlets, but especially because with the fast renewal of the products, there exists a close relationship between the manners of consuming them and manners of producing them. With the expression "to include work to transform it", one could add "to understand the modes of relations to the products to transform work" and to understand the relations with the products to act on the experiment.

T6'(reintroduction of the translated result in the same software application, Systran): Une caractéristique de ces diverses approches de l'utilisability consiste en prolongeant le champ de l'application de l'ergonomie aux produits et aux technologies interactives de la vie quotidienne. Non seulement parce que ceci constitue un nouveau champ des débouchés, mais particulièrement parce qu'avec le renouvellement rapide des produits, là existe une relation étroite entre les façons de les consommer et les façons de les produire. Avec l'expression "pour inclure le travail pour le transformer", on a pu s'ajouter "pour comprendre les modes des relations aux produits pour transformer le travail" et pour comprendre les relations avec les produits pour agir sur l'expérience.

For the evaluation of quality of translation provided by the software applications we identified and classified the errors according to SAE J2450 translation quality standard and QA model and we associated to each found error the corresponding weight according to this metric. In the process of observation, classification and comparative analysis of the quality of translations performed by the two online translation memory software applications, we employed the current version of GT

and Systran, that of June 2016. After having gathered data, we analyzed them using the method for error calculation established by SAE J2450 QA model. The total score of each translation counts the errors of each category found in the translation of each paragraph by the two software applications, while the overall paragraph score is calculated by dividing the sum of numerical weights of the totality of errors committed in the paragraph by the total number of words in the paragraph excerpt from the source-text. We then compared the results obtained following the translations performed by the two online memory translators. We also applied the test of concordance so as to analyze the degree of accuracy of translation as the result of initial translation (from T1 in English to T2 in French/from T4 in French to T5 in English) was reintroduced in the same software application (we retranslated T2 in the same language as that of the source-text T1) and we compared T1 with T3 and T4 with T6, using the same evaluation metric for translation quality.

Figure 1 shows us that when the source-text is in English, a higher degree of accuracy in translation is achieved by the translation memory program GT. This is also true for the test of concordance, when in the second phase, we introduce the text T2 in the same software application (GT) and we retranslate it into English. Comparing T1 with T3 and T1 with T3', we notice a higher degree of concordance between T1 and T3 (GT) than between T1 and T3' (Systran). OPS will be used as the output of the fuzzy models in the next section.
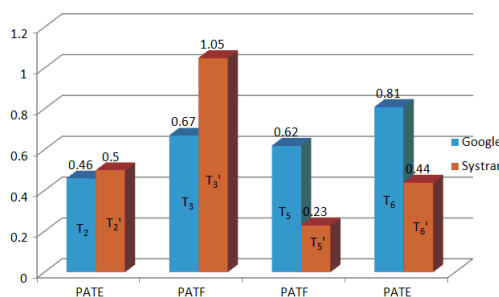

Figure 1

Comparison of Overall Paragraph Scores (OPS), where: PATE – the paragraph of academic text in English and PATF – the paragraph of academic text in French

When the source-language is French, we obtain opposite results: a higher degree of accuracy is achieved by the translation memory software application Systran. The same result was reached when we applied the concordance test to compare T4 with T6 and T6'.

If we make a comparison of scores associated to wrong terms, we draw the following conclusions: when the source-paragraph is in English, we encounter many more errors committed by Systran than by GT. When the source-paragraph is in French, there are more errors committed by GT than by Systran (as shown in Figure 2). WT is the first input to the fuzzy models developed in the next section.
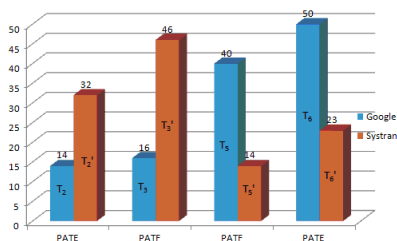
Figure 2
Comparison of Wrong Term (WT) scores

When we reintroduce the result of the paragraph translated from English to French in the same software application so as to compare the new result with the source-text (T1 in English), we remark fewer wrong terms with GT than with Systran. When we reintroduce the result of the paragraph translated from French to English in the same software application so as to compare the new result with the source-text (T4 in French), we note less wrong terms with Systran than with GT.

As regards the comparison related to the number of syntactic errors, when the source-paragraph is in English, Systran commits fewer errors than GT (according to Figure 3). When we reintroduce the paragraph T2 in the same MT application (GT) and T2' in the same software application (Systran) and we retranslate them into English, we obtain a higher accuracy with GT than with Systran. SE is the second input to the fuzzy models developed in the next section.
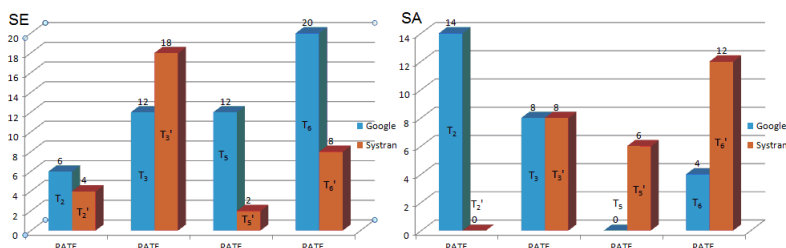

Figure 3
Comparison of Syntactic Error (SE) and Word Structure and Agreement error (SA) scores

When the source-paragraph is in English, the scores measuring the weight of SA errors are higher with GT than with Systran. For the test of concordance, when we reintroduce the result of the paragraph translated from English to French in the same software application in order to compare the new result with the source-text (T1 in English), we get similar scores registering the weight of morphological errors for GT and for Systran.

When the source-paragraph is in French, the scores measuring the weight of SA errors are higher with Systran than with GT. For the concordance test, when we reintroduce the result of the paragraph translated from French to English in the same automated translation application in order to compare the new result with the

source-text (T4 in French), we note higher scores for SA errors with Systran than with GT. The obvious result is that while we translate from English to French, we achieve a higher precision with GT and that, when we translate from French to English, a higher degree of accuracy is reached by Systran (as illustrated in Figure 3). SA is the third input to the fuzzy models developed in the next section.

The weakness of automated translation provided by Systran consists in the fact that it records more wrong terms than GT, inclusively at the test of concordance. By comparison with Systran, GT counts higher scores while it weighs the burden of syntactic errors.

The category of WT is the most important and in the translation performed by Systran, the existence of a higher score recorded for WT is sometimes due to the lack of entry in the dictionary of the automated translator (this is the case of the translation of the English term "utilisability" into French: the automated translator copies the English term as it is while it provides the translation into French). Syntactic errors represent a second major category of errors found with the two software applications whose presence might be justified by the gap between two languages of different roots: Latin root and Germanic root. As regards SA errors, with the two automated translators there is a weakness in the recognition of (noun) determiners and a lack of correspondence and agreement between tenses and verbal moods, a lack of sequence of tenses, as well as erroneous translations of moods and tenses (for example, in the translation T2: "nous ne recentrer nos efforts " and "nous aurons gagné la bataille, mais a perdu la campagne").

As far as the category ME is concerned, a mistake of this kind is characterized as such when it does not befit the other six types of error acknowledged by the metric and the mistake which is most often included in the seventh category is either an awkward choice of preposition or a preposition that misses in the target language. After the comparative study of the behavior of automated translators, we encounter miscellaneous mistakes in translations performed by both GT and Systran, while we can also reveal the fact that GT counts a higher omission error score, as in the translation from English to French the online translation memory application omits terms it does not know. On the other hand, the hybrid translation engine of Systran records a higher error score whilst performing the translation of abbreviations (for example, abbreviations of EU units or institutions). Punctuation errors are common with both automated engines, yet GT repeats the common mistake of not adapting the English inverted commas quotation marks to "guillemets" in the translation from English to French; it is also highly possible that with GT a complex sentence begins with a lower case letter instead of an upper case letter. While the statistical engine favors the choice of synthetic or Saxon Genitive, the hybrid machine translator prefers to choose the analytical or periphrastic Genitive in the translation of texts from French to English. If both engines record mistakes in the translation of idioms, the hybrid application achieves a better performance in the translation of idiomatic expressions from French to English than the statistical translation engine, as it also registers a better

choice of tenses and moods after certain conjunctions (of subordination, such as "bien que" which compulsorily requires a Subjunctive and not an Indicative). Last but not least, we should emphasize that GT is designed to provide a higher translation speed in both translations (from English to French and from French to English) as compared to the hybrid engine Systran which is not only a bit slower, but it also sometimes gets blocked.

3 Evolving Fuzzy Systems Modeling

The flowchart of the online identification algorithm that produces evolving fuzzy systems models is presented in Figure 4, where k is the data sample index, $\mathbf{p}_k$ is the current data point:

$$\mathbf{p}_k = [p_k^1 \quad p_k^2 \dots p_k^{n+1}]^T, \quad \mathbf{p} = [\mathbf{z}^T \quad y]^T = [z_1 \quad z_2 \dots z_n \quad y]^T$$
$$= [p^1 \quad p^2 \dots p^n \quad p^{n+1}]^T \in \mathfrak{R}^{n+1}, \tag{1}$$
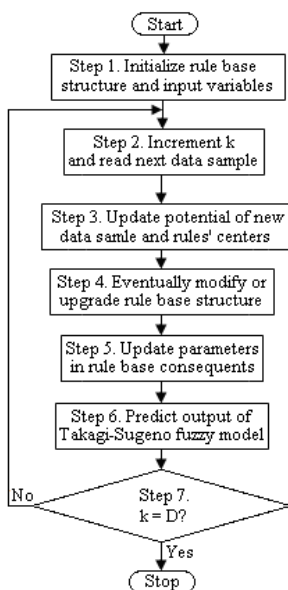

Figure 4
Flowchart of online identification algorithm [6]

T stands for matrix transposition, and the input-output data set is:

$$\{\mathbf{p}_k \mid k = 1 \dots D\} \subset \mathfrak{R}^{n+1}, \tag{2}$$

where D is the number of input-output data points, also called data points or (data) samples.

The rule base of fuzzy models of Takagi-Sugeno type is:

Rule i : IF z_1 IS LT_{i1} AND ... AND z_n IS LT_{in}

THEN $y_i = a_{i0} + a_{i1} z_1 + ... + a_{in} z_n, i = 1...n_R,$ (3)

where $z_j, \ j = 1...n,$ are the input variables, n is the number of input variables, $LT_{ij}, i = 1...n_R, j = 1...n,$ are the input linguistic terms, y_i is the output of the local model in the rule consequent of the rule index $i, i = 1...n_R,$ n_R is the number of rules, and $a_{il}, i = 1...n_R, l = 0...n,$ are the parameters in the rule consequents. The fuzzy model structure includes the algebraic product t-norm as an AND operator and the weighted average defuzzification method.

As specified in the previous sections, the fuzzy models developed in this paper are characterized by a number of $n = 7$ input variables, which are $z_1 = \text{WT}$, $z_2 = \text{SE}$, $z_3 = \text{OM}$, $z_4 = \text{SA}$, $z_5 = \text{SP}$, $z_6 = \text{PE}$ and $z_7 = \text{ME}$, and the output variable is $y = \text{OPS}$. The system inputs are presented in Figure 5, which illustrates the input data for both training and validation (testing).
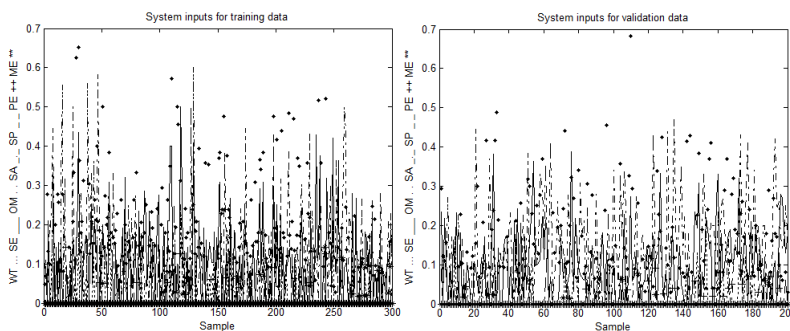

Figure 5

System inputs versus data sample for training data and validation (testing) data

The root mean square error (RMSE) is used as a global performance index in order to compare the four fuzzy models. Its definition is:

$$\text{RMSE} = \sqrt{(1/D)\sum_{k=1}^{D}(y_k - y_{d,k})^2},$$ (4)

where y_k is the model output and $y_{d,k}$ is the real system output, i.e., the actual value of OPS obtained by the application of GT or Systran in the translation from English to French.

The fuzzy model 1, which corresponds to GT using RLS, has evolved to $n_R = 6$ rules, it consists of 132 parameters and its performance on the validation data is $\text{RMSE} = 0.25924$. The fuzzy model 2, which corresponds to GT using wRLS, has

evolved to $n_R = 6$ rules, it consists of 132 parameters and its performance on the validation data is $\text{RMSE} = 0.25923$. The fuzzy model 3, which corresponds to Systran using RLS, has evolved to $n_R = 7$ rules, it consists of 154 parameters and its performance on the validation data is $\text{RMSE} = 0.29652$. The fuzzy model 4, which corresponds to Systran using wRLS, has evolved to $n_R = 7$ rules, it consists of 154 parameters and its performance on the validation data is $\text{RMSE} = 0.29653$.

These results show that for the data set considered in this paper the best performance is obtained by the fuzzy models that model the GT-based translation, namely the fuzzy models 1 and 2. The use of either the RLS algorithm or the wRLS the algorithm to update the rule consequents of the fuzzy models in the online identification algorithm does not affect the results.

The outputs of the fuzzy model 2 and of the real system output are illustrated in Figure 6 for both training data and validation (testing) data.
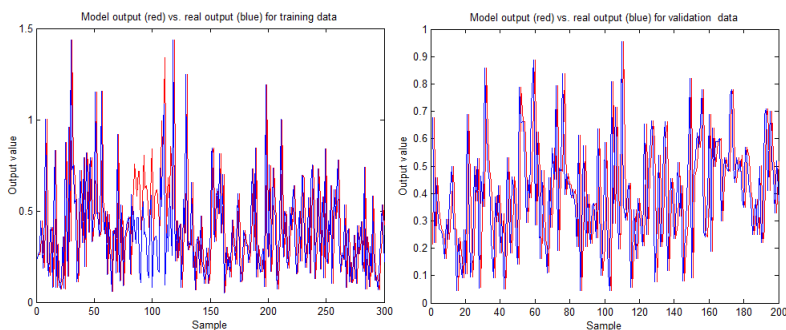

Figure 6

System output (OPS) versus data sample of evolving fuzzy model and real system for training data and validation (testing) data

The results presented in Figure 6 are encouraging. However, different conclusions are expected to be obtained for other applications [33-40].

Conclusions

This paper has proposed evolving fuzzy models, which give the Overall Paragraph Score in translation using these seven categories of primary errors as inputs. Four fuzzy models have been developed using an incremental online identification algorithm applied to the translation from English to French. The data sets have been obtained by processing the results obtained by the application of GT and Systran.

By using the excerpts from academic texts and two source-languages (English and French), it is not easy to discern that on the occasion of evaluation of translation provided by the hybrid translation software application, carried out with the help of SAE J2450 standard, we record a higher score than that of the evaluation of

translation made by the statistical software application for all seven categories of errors. It is true that statistical software performs better than the hybrid software in the translation of terms, but it is also obvious that in the translation from French to English and on applying the test of concordance between the original paragraph in French and the translation resulted from the reintroduction of the result translated from French to English in the same software application (Systran), we obtain a higher degree of accuracy and precision for the hybrid software than for the statistical software.

All these operations of translation and the passages from English to French and from French to English in translation performed by automated engines, as well as the process of evaluation of translation quality with the help of SAE J2450 metric, of the test of concordance and of the assessment made with the statistics software Excel, provide a reference guide for automated translation analysis which could prove its utility and efficiency for students in translation and for professional translators, since it gives them the chance to enhance their analytic thinking skills and to expand their researches on the existing differences between degrees of accuracy of translation engines chosen in their work, on the weak and strong points of each of these automated translators, on the recurrent errors of each software application, all the more so as the error is always found in the center of their perception of translation learning.

The future research will be focused on the development of neural network models and other appropriate applications with various optimization approaches [41-50] included in the identification algorithm. The development of fuzzy models for the translation from French to English will require modifications in the structure of the identification algorithms.

Acknowledgement

References

[1]   R.-E. Precup, H.-I. Filip, M.-B. Radac, C. Pozna, C.-A. Dragos, S. Preitl: Experimental Results of Evolving Takagi-Sugeno Fuzzy Models for a Nonlinear Benchmark, Proceedings of 2012 IEEE 3rd International Conference on Cognitive Infocommunications, Kosice, Slovakia, 2012, pp. 567-572

[2]   R.-E. Precup, H.-I. Filip, M.-B. Radac, E. M. Petriu, S. Preitl, C.-A. Dragos: Online Identification of Evolving Takagi-Sugeno-Kang Fuzzy Models for Crane Systems, Applied Soft Computing, Vol. 24, 2014, pp. 1155-1163

[3]   R.-E. Precup, E.-I. Voisan, E. M. Petriu, M.-B. Radac, L.-O. Fedorovici: Implementation of Evolving Fuzzy models of a Nonlinear Process, Proceedings of 12th International Conference on Informatics in Control, Automation and Robotics, Colmar, France, 2015, Vol. 1, pp. 5-14

[4]   R.-E. Precup, E.-I. Voisan, E. M. Petriu, M.-B. Radac, L.-O. Fedorovici: Gravitational Search Algorithm-Based Evolving Fuzzy Models of a Nonlinear Process, Informatics in Control, Automation and Robotics, J. Filipe, K. Madani, O. Gusikhin, and J. Sasiadek, Eds., Lecture Notes in Electrical Engineering, Springer International Publishing, Cham, Vol. 383, 2016, pp. 51-62

[5]   R.-E. Precup, T.-A. Teban, T. E. Alves de Oliveira, E. M. Petriu: Evolving Fuzzy Models for Myoelectric-Based Control of a Prosthetic Hand, Proceedings of 2016 IEEE International Conference on Fuzzy Systems, Vancouver, Canada, 2016, pp. 72-77

[6]   R.-E. Precup, M.-B. Radac, E. M. Petriu, R.-C. Roman, T.-A. Teban, A.-I. Szedlak-Stinean: Evolving Fuzzy Models for the Position Control of Twin Rotor Aerodynamic Systems, Proceedings of 2016 IEEE 14th International Conference on Industrial Informatics, Poitiers, France, 2016, pp. 237-242

[7]   J. V. Ramos, A. Dourado: On Line Interpretability by Rule Base Simplification and Reduction, Proceedings of European Symposium on Intelligent Technologies, Hybrid Systems and Their Implementation on Smart Adaptive Systems EUNITE 2004, Aachen, Germany, 2004, pp. 1-6

[8]   L. Aires, J. Araújo, A. Dourado: Industrial Monitoring by Evolving Fuzzy Systems, Proceedings of Joint 2009 IFSA World Congress and 2009 EUSFLAT Conference, Lisbon, Portugal, 2009, pp. 1358-1363

[9]   A. Bantaş, E. Croitoru: Didactica Traducerii, Editura Teora, Bucharest, 1999 (in Romanian)

[10]   D. Toudic, K. Hernandez Morin, F. Moreau, F. Barbin, G. Phuey: Du Contexte Didactique aux Pratiques Professionnelles: Proposition d'une Grille Multicritères pour l'Évaluation de la Qualité en Traduction Spécialisée, ILCEA Online, No. 19, 2014, pp. 1-19

[11]   L. Bowker, M. Ehgoetz: Exploring User Acceptance of Machine Translation Output: A Recipient Evaluation, Across Boundaries: International Perspectives on Translation, D. Kenny, K. Ryou, Eds., Cambridge Scholars Publishing, Newcastle upon Tyne, 2007, pp. 209-224

[12]   S. Green, J. Heer, C. D. Manning: The Efficacy of Human Post-Editing for Language Translation, Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, Paris, France, 2013, pp. 439-448

[13]   P. Koehn: 20 Years of Statistical Machine Translation, University of Edinburgh Press, Edinburgh, 2009

[14]   Y. Bar-Hillel: The Present Status of Automatic Translation of Languages, Advances in Computers, No. 1, 1960, pp. 91-163

[15]   R. Ryan: Traduction Technique: les Langues Contrôlées au Service de l'Ergonomie Documentaire, ILCEA Online, No. 14, 2011, pp. 1-18

[16]   A. Hartley, A. Popescu-Belis: Évaluation des Systèmes de Traduction automatique, in Évaluation des Systèmes de Traitement de l'Information, S. Chaudiron, Ed., Éditions Hermès, 2004, pp. 311-335

[17]   S. Nirenburg: Knowledge-Based Machine Translation: Parameter Estimation, Machine Translation, No. 4, 1989, pp. 5-24

[18]   M. Costa-Jussà, M. Farrús, J. Mariñom, J. Fonollosa: Study and Comparison of Rule-Based and Statistical Catalan-Spanish Machine Translation Systems, Computing and Informatics, Vol. 31, 2012, pp. 245-270

[19]   P. Brown, S. Della Pietra, V. Della Pietra, R. Mercer: The Mathematics of Statistical Machine Translation: Parameter Estimation, Computational Linguistics, Vol. 19, No. 2, 1991, pp. 263-311

[20]   M. Nagao: A Framework of a Mechanical Translation between Japanese and English by Analogy Principle, Artificial and Human Intelligence, R. Banerji, Ed., Elsevier Science Publishers, 1984, pp. 173-180

[21]   D. Turcato, P. McFetridge, F. Popowich, J. Toole: A Unified Example-Based and Lexicalist Approach to Machine Translation, Proceedings of 8th International Conference on Theoretical and Methodological Issues in Machine Translation, Chester, UK, 1999, pp. 33-43

[22]   A. Boretz: AppTech Launches Hybrid Machine Translation Software, SpeechTechMag.com published on the 2nd of March 2009, http://www.speechtechmag.com/Articles/News/Speech-Technology-News-Features/AppTek-Launches-Hybrid-Machine-Translation-Software-52871.aspx

[23]   S. Hunsicker, C. Yu, C. Federmann: Machine Learning for Hybrid Machine Translation, Proceedings of 7th Workshop on Statistical Machine Translation, Association for Computational Linguistics, Montreal, Canada, 2012, pp. 312-316

[24]   V. Drexler: "I Don't Translate, I Create!" - An On-line Survey on Uniformity versus Creativity in Professional Translations, Anchor Academic Publishing, Hamburg, 2016

[25]   R. Rubino: Traduction Automatique Statistique et Adaptation à un Domaine Spécialisé, PhD Thesis, University of Avignon and the Vaucluse, Montreal, Canada, 2011

[26]   S. Green, S. Wang, J. Chuang, J. Heer, C. Manning: Human Effort and Machine Learnability in Computer-Aided Translation, EMNLP, 2014, pp. 1225-1236

[27]   M. W. Madsen: The Limits of Machine Translation, MSc Thesis, University of Copenhagen, Copenhagen, Denmark, 2009

[28]   K. Papineni, S. Roukos, D. Ward: Corpus-based Comprehensive and Diagnostic MT Evaluation: Initial Arabic, Chinese, French and Spanish Results, Proceedings of HLT, Second International Conference on Human Language Technology Research, M. Markus, Ed., Morgan Kaufmann, 2002, pp. 132-137

[29]   SAE: Surface Vehicle Recommended Practice, http://www.apex-translations.com/documents/sae_j2450.pdf, 2001

[30]   A. Secară: Translation Evaluation – A State of the Art Survey, Proceedings of the EColorMe/Mellange Workshop, Leeds, UK, 2005, pp. 39-44

[31]   S. M. Cowan: Information Literacy: The Battle We Won That We Lost?, Portal: Libraries and the academy, Vol. 14, No. 1, 2012, pp. 23-32

[32]   J. Barcenilla, J.-M.-C. Bastien: L'Acceptabilité des Nouvelles Technologies: Qquelles Relations avec l'Ergonomie, l'Utilisabilité et l'Expérience Utilisateur ?, Le Travail Humain, Vol. 72, No. 4, 2009, pp. 311-331

[33]   R.-E. Precup, S. Preitl, M. Balas, V. Balas: Fuzzy Controllers for Tire Slip Control in Anti-lock Braking Systems, Proceedings of IEEE International Conference on Fuzzy Systems, Budapest, Hungary, 2004, Vol. 3, pp. 1317-1322

[34]   R.-E. Precup, S. Preitl: PI-Fuzzy Controllers for Integral Plants to Ensure Robust Stability, Information Sciences, Vol. No. 20, 2007, pp. 4410-4429

[35]   R.-E. Precup, M.-B. Radac, M. L. Tomescu, E. M. Petriu, S. Preitl: Stable and Convergent Iterative Feedback Tuning of Fuzzy Controllers for Discrete-Time SISO Systems, Expert Systems with Applications, Vol. 40, No. 1, 2013, pp. 188-199

[36]   A. L. Kazakov, A. A. Lempert: On mathematical Models for Optimization Problem of Logistics Infrastructure, International Journal of Artificial Intelligence, Vol. 13, No. 1, 2015, pp. 200-210

[37]   Á. Takács, D. Á. Nagy, I. J. Rudas, T. Haidegger: Origins of Surgical Robotics: From Space to the Operating Room, Acta Polytechnica Hungarica, Vol. 13, No. 1, 2016, pp. 13-30

[38]   I. J. Rudas, J. Gáti, A. Szakál, K. Némethy: From the Smart Hands to Tele-Operations, Acta Polytechnica Hungarica, Vol. 13, No. 1, 2016, pp. 43-60

[39]   B. Takács, R. Dóczi, B. Sütő, J. Kalló, T. A. Várkonyi, T. Haidegger, M. Kozlovszky: Extending AUV Response Robot Capabilities to Solve Standardized Test Methods, Acta Polytechnica Hungarica, Vol. 13, No. 1, 2016, pp. 157-170

[40]   S. Anbarasi, S. Muralidharan: Enhancing the Transient Performances and Stability of AVR System with BFOA Tuned PID Controller, Control Engineering and Applied Informatics, Vol. 18, No. 1, 2016, pp. 20-29

[41]   I. Škrjanc, S. Blažič, O. E. Agamennoni: Identification of Dynamical Systems with a Robust Interval Fuzzy Model, Automatica, Vol. 41, No. 2, 2005, pp. 327-332

[42]   F. G. Filip: Decision Support and Control for Large-scale Complex Systems, Annual Reviews in Control, Vol. 32, No. 1, 2008, pp. 61-70

[43]   A. Gajate, R. E. Haber, P. I. Vega, J. R. Alique: A Transductive Neuro-fuzzy Controller: Application to a Drilling Process, IEEE Transactions on Neural Networks, Vol. 21, No. 7, 2010, pp. 1158-1167

[44]   J. Vaščák, K. Hirota: Integrated Decision-Making System for Robot Soccer, Journal of Advanced Computational Intelligence and Intelligent Informatics, Vol. 15, No. 2, 2011, pp. 156-163

[45]   Zs. Cs. Johanyák: Fuzzy Modeling of Thermoplastic Composites' Melt Volume Rate, Computing and Informatics, Vol. 32, No. 4, 2013, pp. 845-857

[46]   P. Baranyi, Y. Yam, P. Varlaki: TP Model Transformation in Polytopic Model-Based Control, Taylor & Francis, Boca Raton, FL, 2013

[47]   O. Castillo, E. Lizárraga, J. Soria, P. Melin, F. Valdez: New Approach Using Ant Colony Optimization with Ant Set Partition for Fuzzy Control Design Applied to the Ball and Beam System, Information Sciences, Vol. 294, 2015, pp. 203-215

[48]   O. Arsene, I. Dumitrache, I. Mihu: Expert System for Medicine Diagnosis Using Software Agents, Expert Systems with Applications, Vol. 42, No. 4, 2015, pp. 1825-1834

[49]   A. Khmelev, Y. Kochetov: A Hybrid Local Search for the Split Delivery Vehicle Routing Problem, International Journal of Artificial Intelligence, Vol. 13, no. 1, 2015, pp. 147-164

[50]   L. Zhu, C.-X. Fan, Z.-G. Wen, H.-R. Wu: Coverage Optimization Strategy for WSN based on Energy-aware, International Journal of Computers, Communication and Control, Vol. 11, No. 6, 2016, pp. 877-888