*International Journal of Computational Linguistics and Applications*, Vol. 6, No. 2, 2015, pp. 43–59
Received 01/02/2015, Accepted 27/02/2015, Final 07/04/2015.
ISSN 0976-0962, http://ijcla.bahripublications.com

# A New Machine Translation Decoder Based on Artificial Immune System

MANEL AMMAR
SALMA JAMOUSSI
*Sfax University, Tunisia*

ABSTRACT

*This paper focuses on decoding as main part of statistical machine translation. Decoding is considering as a NP-complete algorithm that requires intelligent heuristics to and optimum solutions. In order to solve this problem, we proposed a decoder named DAIS based on the meta-heuristic of artificial immune system. The evaluation is performed on two different corpora. The obtained translations show that the proposed approach obtains encouraging results by comparing them to those of the most known decoders in the field.*

## 1. INTRODUCTION

The main task of machine translation (MT) is to translate from one natural language to other target languages. Reduce human intervention is one of the main objectives in this task. A statistical MT (SMT) presents a promising avenue for eliminating experts role, improving translations quality and reducing costs. To make an SMT system, we need three complementary components. First, a language model which is used to and syntactically correct results in the target language. It is trained from monolingual text corpora. Second, the translation model which needs a parallel corpora in order to align the different source sequences with their probable target translations. Then, it assigns a probability value characterizing their appearances in the training corpora. Finally, a decoder produces an optimal statistical translation among all possible translations

by using intelligent search algorithms. These basic components of the SMT can be likened to the noisy-channel approach [1]. Infact, The main goal of an SMT system is to produce the best target sentence *T** for a given source sentence *S*. It is the decoder that will search for this solution in order to compromise between maximizing the translation probability $P(T|S)$ and the language model probability $P(T)$. *T** is found from the following Bayes theorem:

$$T^* = argmax_T P(T|S) = argmax_T P(T) \times P(T|S) \quad (1)$$

In this paper, we focus on developing a decoder called DAIS, abbreviation for Decoding with Artificial Immune System. DAIS is a SMT decoder based on the meta-heuristic of immune systems. DAIS was experimented on a French English translation task and inversely. It uses two fundamental SMT components namely, translation model (TM) and the language model (LM). the obtained results show that our proposed decoder has better performance than Moses decoder. We also scored an acceptable execution time comparable to the time taken by Moses. In this paper, we have studied state-of-the-art decoding task. Next, we present the meta-heuristic of immune system as well as its application in the decoding task in SMT. Finally, we evaluate its performance by comparing the obtained results with DAIS to other named decoder such as Moses and Google translation.

## 2. RELATED WORKS

The fundamental dilemma of the decoding task is the huge search space. Thus, it is not possible to check all translation possibilities because of the combinatorial explosion of feasible hypotheses. To make the decoding task feasible and effective, the key solution was to use the optimization methods based on intelligent meta-heuristics. These methods are well suited for resolving this type of problems with huge search spaces. We distinguish two broad classes of methods. The first includes local search methods such as beam search, dynamic programming, etc. The second is

the class of global methods as evolutionary methods. Several decoders have been proposed in the literature. The majority of them are based on local optimization methods. This is thanks to their neighbourhood principle allowing the use of a relatively small memory space. We cite in this context, Moses [3] and Marie [4] decoders which realize different implementations of beam search algorithm. We also find in the literature decoders based on dynamic programming [5] and on greedy search [6].

According to our knowledge, the number of decoders implementing bio-inspired methods is very limited. We find only the works of [7] which use genetic algorithm to develop a decoder called GADecoder (Genetic Algorithm Decoder). Since it is a word based decoder, it can find good translations for words but it has difficulties to find the best word ordering. With 4-grams language model, the results were not good enough as Pharaoh [2] and Google Translation. We have noticed that bio-inspired methods are rarely used in SMT field. Indeed, they avoid local optima and provide a global view of the search space thanks to their candidates distributed intelligently and contributed each one to build the final solution.

## 3. FUNDAMENTALS AND MAIN COMPONENTS OF THE ARTIFICIAL IMMUNE SYSTEM (AIS)

**In nature**, the immune system plays a very important role in the survival of vertebrates. Indeed, the defense mechanism aims to protect the body against unauthorized intruders no belonging to self named antigens. An immune system is then faced with the problems of detection, identification and response to pathogens. The first step of an immune system is to determine antigens no belonging to self by the mechanism of negative selection. Then clonal selection mechanism will be trigged in order to prepare the adequate immunity response allowing their elimination. The basic actors in the immune system are the white blood cells or lymphocytes. They produce receptors called antibodies, with which they can discriminate self from non-self and neutralize pathogens elements. Neutralization of an antigen is made by a connection between the antibody and a specific part of the

intruder antigen. More the degree of affinity of this connection is strong, the system is able to eliminate the concerned antigen. There are two main types of natural immunity. The first one is the innate immunity; it is an elementary response against an antigen firstly encountered in the human body. The second is adaptive immunity; it is a stored immune response indicating that the human body has already identified this antigen and the antibody specific for its neutralization is already built and ready to respond.

**In computer science**, the artificial immune system (AIS) is an interesting biological inspiration for solving optimization problems [8]. On a computer view, the negative selection algorithm is based on the idea of generating a set of malicious elements and tests the ability of such system to determine them [9]. Thus, the clonal selection algorithm uses the set generated by the negative selection algorithm in order to nd the best solution to the problem. It aims to produce an initial population of solutions. These solutions will be cloned several times according to their degrees of affinity with the antigen. Clones results will undergo mutations by modifying some of their characteristics in order to attain a higher degree of affinity with the antigen. Algorithm 1 explains the general principle of artificial clonal selection.

```
Algorithm 1: The artificial clonal selection
Input: S= a set of antigens to be recognized
Output: M= a set of best antibodies met\\
  Begin
  Create an initial random set of antibodies, A
  for each Si in S
    Determine the affinity of Si with each antibody in A,
    Generating clones for each antibody in A
    proportionally to its affinity degrees
    Mutate attributes of these clones
    Place a copy of the highest affinity antibodies into
    the memory set, M
```

Figure 1. *The principle of the clonal selection algorithm for the SMT decoding problem*

## 4. AN ARTIFICIAL IMMUNE SYSTEM FOR MACHINE TRANSLATION

AIS meta-heuristic has a number of features that are so interesting to test them for resolving the decoding problem in SMT. Indeed, there are common points between the decoding as a computer problem and the functioning of a real immune system. The latter one must react quickly and correctly to protect the body. For an effective immune response, the human body must have the necessary information on the specifics of each antigen. The same for the decoding problem in SMT, the goal is to acquire characteristics of pair of language in order to produce quickly and efficiently target translations. In addition, the clonal selection technique enables antibodies to multiply proportionally to their quality of neutralization of an antigen; we notice that its implementation will promote the less costly translations. To apply AIS for the SMT task we need first to define some preliminary concepts:

- The antigen in SMT (*AG*): it is the source sentence to be translated expressed in a source language;
- The antibody (*AC*): the intended aim; it is the element allowing neutralization of the antigen. It corresponds to the target sentence expressed in the desired language. A good antibody must be both faithful to the source sentence content and syntactically correct in the target language;
- The innate response is the policy that we have adopted for solving the decoding problem. We considered each source sentence as a new antigen to be neutralized.

## 5. THE NEGATIVE SELECTION ALGORITHM

The negative selection algorithm enables the immune system to discriminate self from non-self cells. Transposing this principle to our decoding task seems very useful. Infact, upon penetration of an antigen that corresponds to a non-self cell, the artificial negative selection is triggered. First, it locates the area in the

search space in which we can find the target solution. So, it generates a set of translation options denoted *D*. This set is composed of translation options which each option is defined by a source fragment, a target fragment, the *TM* value and the number of source words covered by this option. Second, we prepare and reduce the set *D* in order to promote the most promising translation options. We eliminated unattractive options using the pruning technique. It consists in eliminating all translation options having high cost, and this by comparing the TM value of each option to a fixed threshold. The following algorithm 2 describes the mechanism of negative selection adopted in our decoding algorithm.

```
Algorithm 2: Negative selection algorithm in DAIS decoder
Input E= the search space composed of translation options: ei
while (i <Card(E))
  if(ei covers a part of AG )
    if(TM(ei) < thresholdPruning )
      ignore the option ei from the search space
    else
      sort ei in set D taking into account the number of words
      covered in the source sentence and its TM value
      (equiprobable scheduling
```

Figure 2. *Construction of the search space for a given antigen AG*

## 6. CLONAL SELECTION ALGORITHM

The clonal selection aims to produce specific antibodies allowing the neutralization of pathogens. In our case, the clonal selection algorithm is the heart of our translation system. Indeed, we admit that the artificial neutralization consists in finding the best target translation of a given source sentence. The principle of this algorithm illustrated in Figure 3 is based on four operators namely evaluation operators, cloning, mutation and communication. The initialization of our translation system consists in producing a random set of antibodies population. So, we need to use the subspace given by the negative selection algorithm to form a set of initial hypotheses in the target

language that we called initial antibodies. The system thus, needs to improve those hypotheses to find the best antibody.

6.1. *Affinity evaluation*

To evaluate the affinity of selected antibodies, we adopted the same scoring function implemented in Moses decoder [3]. The goal is to find the antibody $T$ (target sentence) for a given antigen $S$ (source sentence), which maximizes $Affinity\ (T, S)$. This score function is presented as follows:

$$Affinity\ (T, S) = \lambda_{tr} \times log(TM(S|T)) + \lambda_{lm} \times log(LM\ (T)) - \lambda_{w} \times log(exp^{w}\ (T)) \quad (2)$$

Where $TM(S|T)$, $LM$(T) and $exp^{w\,(T)}$ represent respectively the translation model probability, language model probability and the penalty on the length of the target sentence $T$. $\lambda_{tr}$, $\lambda_{lm}$ and $\lambda_{w}$ are the weights of the translation model, language model and length penalty. Indeed, a good antibody is one that maximizes the objective function Affinity $(T, S)$. To estimate lambda, we use MERT tuned system from Moses decoder.

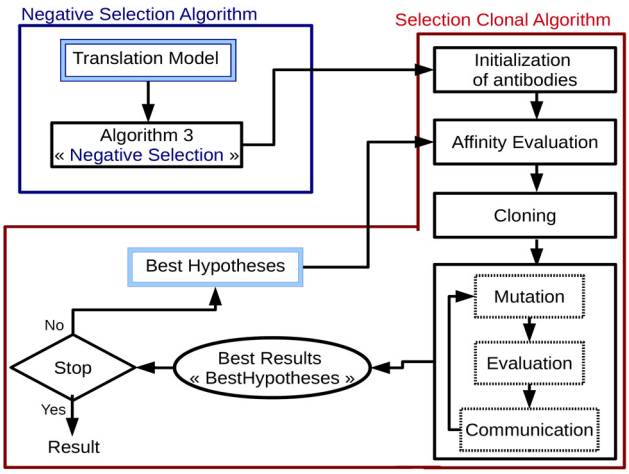


Figure 3. *The principle of the clonal selection algorithm for the SMT decoding problem*

6.2. *Cloning*

The cloning technique is to make identical copies to hypotheses proportionally to their evaluation scores. As a result, the antibody that has the best affinity score has the higher number of clones. The following formula calculates a percentage characterizing number of clones that deserves a given antibody:

$$nbClones_{AG}(AC_i) = \frac{Affinity(AC_i, AG)}{\sum_{j=0}^{NCI} Affinity(AC_j, AG)} \times NCI \quad (3)$$

Where NCI is the number of initial candidates. This parameter had been set empirically.

6.3. *Mutation*

The mutation alters the characteristics of different clones in order to evolve them. We implemented two types of mutation. The first one substitutes an option in an antibody by another from the search space. The second consists in changing the positions of translation options in a given antibody.

**The mutation of translation options**: In a first step and for each clone we fixed the number of mutations ($NM(AC_i)$) to the number of translation options presented in $AC_i$. Each iteration of mutation begins by randomly choosing the translation option to be mutated ($OP_j$). Then, we specify the number of mutations that merit $OP_j$. Thus, a translation option $OP_j$ having a low a affinity requires more mutation than another option having improved a affinity. The following formula seeks the number of mutations required for a translation option $nbMut(OP_j)$ in an antibody $AC_i$:

$$nbMut(OP_j) = (1 - \frac{TM(OP_j)}{\sum_{k=0}^{NM(AC_i)} TM(OP_k) + \varepsilon}) \times NM(AC_i) \quad (4)$$

The second step consists in consulting the search space to select the new translation option that will replace the old one. So, we used the method of wheel selection explained in [7]. Indeed, this

algorithm is applicable for ordered search space that corresponds well to our case.

The mutation of positions of translation options: Translation from source into target language requires often changing positions of words in the target sentence to be syntactically correct in the target language. In our decoder, a translation option can interchanges its position only with options that directly precede or follow it. A correct inversion of position words is useful information to communicate to other antibodies. This communication is elaborated through an enrichment of the search space by a new translation option in order to take advantage for the next clones.

At antibody level, a good mutation signifies that the latter was able to improve it's affinity score. So it can be very close to neutralize the source sentence. For this, we gave him a chance to make another mutation ($NM(\mathrm{AC}_i) = NM(AC_i) + 1$). On the contrary, an antibody which makes a deteriorating mutation of it's a affinity score will be penalized by reducing its chance of one mutation ($NM(AC_i) = NM(ACi) - 1$).

6.4. *Communication between clones*

For bio-inspired algorithms, communication between candidates represents the ideal solution to cope with the combinatorial explosion and improves the convergence speed and quality of AIS algorithm.

**Communication adopted by the AIS meta-heuristic**: According to Figure 3, the AIS algorithm uses antibodies results of iteration $i$ to establish the iteration $i + 1$. These antibodies are not only the initial candidates for the new iteration, but also a new search space for this iteration. Thus, to improve the quality of solutions, two possibilities can be envisaged: either we increase the number of antibodies participating in the search for a solution. As a result, we obtain a rich search space in each iteration. Or we try to improve the quality of antibodies in iteration $i$, so the algorithm in this case, requires only a few iterations to converge. Due to limited capacities of our machines,

we chose the second possibility and this by proposing a new type of communication for AIS algorithms.

**Proposed communication**: The proposed technique consists in taking advantage of the information provided by an antibody when it mutes one of its translation options or change its positions. The affinity score allows us to measure the impact of this mutation on the translation quality. Thus, if the score has improved after a replacement of an option $OP_{anc}$ by other $OP_{new}$, we propose then, to record a slight improvement (bonus) in the TM value of $OP_{new}$ using the following formula:

$$TM_{OP_{new}} = TM_{OP_{new}} + (TM_{OP_{new}} \ bonus) \quad (5)$$

Conversely, if an option does not improve the quality of the translation, then it undergoes a reduction to its TM value ($malus$) as follows:

$$TM_{OP_{new}} = TM_{OP_{new}} - (TM_{OP_{new}} \times malus) \quad (6)$$

Note that the couple ($bonus$, $malus$) is determined empirically. So, clones have right to change the structure of the search space (fostering option over another).

6.5. *Stop condition*

The evaluation process, cloning and mutation are repeated until a finite number of iterations. In natural immune system, one mutation operation creates a new antibody different from original which in turn is susceptible to undergo a new cloning iteration. The application of AIS algorithm in this way will cause an exponentially higher complexity. This is caused by the cloning operator allowing an exponential increase in the number of candidates. To avoid this problem and make possible the application of AIS algorithm, we conducted a series of mutations on each antibody before starting a new cloning iteration (Figure 3). The number of iterations taken by DAIS to converge to a good solution is estimated as follow:

$$NbItr\ (AC_i) \simeq NbClones(AC_i)\ \text{X NM(ACi)}\ \text{X cts} \quad (7)$$

We can conclude that $nbItr$ is a dynamic parameter that is influenced by the length of the source sentence, the number of initial candidates ($NCI$) and the constant ($cts$ is a constant indicating the number of times that cloning process was repeated). A good configuration can give better results in a reasonable time. For example, for a short source sentence, we have set a very low $cts$ to avoid a long and unnecessary calculation. For a long source sentence, it would be effective to increase the number of antibodies $NCI$ and $cts$ to explore more feasible solutions before converging.

The general algorithm of the DAIS decoder is illustrated in algorithm 3.

## 7. Results and Discussion

### 7.1. *Corpora*

An SMT process requires two basic steps: learning and testing. Each one requires bilingual and aligned corpora. Both training and test corpora must be from the same domain. DAIS decoder has been tested on two different corpora: French and English sentences. First, we used the English French "WIT[3]" [10] parallel corpora extracted from TED Talks. Second, we test our translation system on a French English bilingual corpora that is "OpenOffice[1]" containing a collection of documents describing the functions offered by the open office tools. Table 1 summarizes the characteristics of these two corpora. In the training phase, the translation model is obtained using the GIZA++ tool [11]. It allows the alignment of bilingual corpora in order to extract sub-translations with their probabilities. The language model is learned respectively on the French and English monolingual corpora. It is used to assign a probability value to a word sequence to qualify the correct construction of this latter in

[1] http://opus.lingfil.uu.se/OpenOffice.php

the target language. DAIS decoder uses a tri-gram language model generated by SRILM tool [12].

### 7.2. *Evaluation of DAIS decoder*

The adopted translation system requires the adjustment of a number of parameters. So, we make several attempts to detect the best configuration. The best parameter values found are summarized in Table 2.

```
 Algorithm 3: General algorithm for DAIS decoder
Input: AG = antigen (sentence to be translated)
--Application of negative selection algorithm--
  construction of search space of AG according to the algorithm 2
--Application of clonal selection algorithm--
  initialization of NCI candidates
  repeat
    while(i <NCI)
    Evaluation: score= Affinity(ACi, AG) (Equation 2)
    Cloning: nbAC= clonage(ACi) (Equation 3)
    while(j <nbAC)
      NM(ACij) = Card(ACij)= number of translation options in ACij
      while(k <NM(ACij))
      Mutation: Randomly select an option to be mutated: OP
      nbOP: number of mutations that deserves OP (Equation 4)
      while (r <nbOP)
         Substitute OP with another from the search space using
         wheel selection algorithm
         Evaluation: score' after this mutation
         if(score' > score)
            Assign a bonus to OP (Equation 5)
            Another chance to mutation given for (ACij)
        else
            Mutate the position of OP with its neighbors: OPneigh
            Evaluation: score'' after this mutation of position
            If (score'' > score')
                Add the new OPneigh to the search space
            else
              Assign a malus to OP (Equation 6)
              Reduce the number of mutation given for (ACij)
    until(cts): number of iterations was satisfied.
 return ACmax: the best antibody met
```

Figure 4. *Algorithm 3: General algorithm for DAIS decoder*

The evaluation is based on two main criteria, the quality of the translation reflected by score BLEU [13] and score TER [14], and the execution time. We compared our DAIS decoder with other reference decoders such as Moses and Google translator. Table 3 shows results obtained with the test corpora "test2012" by the different translators (DAIS, Moses and Google translation). We can notice that the BLEU values obtained by DAIS are better than those obtained by Moses. For example, for 4-grams, our DAIS decoder gets a BLEU value equals to 19:29 against 13:71 obtained with Moses decoder. Results also showed that Google translator generates the better quality translations than DAIS and Moses with a BLEU score equal to 42:12.

For the TER score, we found 0:65 translation error rate for DAIS decoder in test2012. We are better than Moses which has 0:68. Google is the best with TER equal to 0:46. Note that TER is an error metric for machine translation that measures the number of edits required to change a system output into one of the references.

Table 1. *Characteristics of the used corpora*

| Corpora | WIT[3] | | OpenOffice | |
|---|---|---|---|---|
| | English-French | | French-English | |
| Type | Train | Test2012 | Train | Test |
| Number of sentences | 186510 | 1124 | 24500 | 1000 |

Table 2. *The best parameter values found after several execution attempts*

| **Parameters** | **test2012** | **OpenOffice** |
|---|---|---|
| nbItr | 1000 | 200 |
| ($\lambda_{tr}$ , $\lambda_{lm}$, $\lambda_{w}$) | (3 , 2 , 0.7) | (1 , 2 , -1) |
| NCI | 50 | 20 |
| (bonus, malus) | (0.1 , 0) | (0.1 , 0) |

In Table 3, we note that DAIS required an execution time higher than Moses decoder (an average of 1 minutes per sentence with DAIS against 0:5 minutes with Moses). Google has the most efficient time translation with a mean of 0:02 seconds per sentence.

Indeed, we share the same learning base with Moses (the translation model and the language model). However, Google translator uses an efficient language model that covers almost all the linguistic peculiarities of the French language and a complete translation model.

Also, to ensure that DAIS decoder is independent of the pair of used languages, we test it with a French English corpora which is "Open Office". We used a sample of 1000 French sentences sources to compare the behavior of the different translators. From Table 4, we note that the results clearly confirm our previous conclusions. The obtained results are excellent even exceeding those given by Google with a gap equal at 3:4. Furthermore, we note that our DAIS decoder leaving an acceptable execution time equal to 12 seconds per sentence. However, it is higher than the time taken by Moses (3 seconds per sentence) Google (0:002 seconds per sentence). Same for the TER score, DAIS is the best with TER = 0:46, the second is google $TER$ = 0:69 and the third is Moses TER = 0:89. We note that the obtained results with OpenOffice corpora are better than those obtained with WIT[3]. This implies that our decoder performs better from French to English than vice versa.

Table 3. *Comparison of translation results in terms of BLEU score and execution time*

| **nGram** | **Google** | **Moses** | **DAIS** |
|---|---|---|---|
| **4-grams** | 42.12 | 13.71 | 19.29 |
| **Time: min/ph** | 0.02 | 0.5 | 1 |

Table 4. *Comparison of translation results in terms of BLEU score and execution time obtained by DAIS and Google with the test corpora of "Open Office"*

| **Comparison characteristics** | **Moses** | **Google** | **DAIS** |
|---|---|---|---|
| **4-grams** | 10.72 | 22.4 | 25.8 |
| **Execution Time: sec/ph** | 3 | 0.002 | 12 |

## 8. CONCLUSIONS AND PERSPECTIVES

The decoding is a main part of statistical machine translation. It aims to find the right combinations of target fragments by using a decoding algorithm. The optimization method must manage a difficult compromise between the quality of translations and the execution time. In this context, we used the artificial immune system algorithm inspired by the biological immune system. Indeed, we have developed a SMT decoder named DAIS that enables each candidate to build its own solution and communicate with others candidates in order to form cooperatively the best solution for the problem. We trained and tested our decoder system on two corpora; "WIT3" and "Open Office". To evaluate the performance of DAIS, we compared it to other known decoders as Moses, which is based on beam search algorithm and the Google translator that uses very large databases. The comparison is based on translation quality measured by BLEU score and execution time. We have noticed that DAIS decoder gives better performance than Moses. Results are generally quite encouraging, they are comparable to those obtained by Moses but they are still further than those obtained by the Google translator thanks to the richness of its database.

In future work we will focus on the concept adaptive reaction. It seems more efficient and especially when we increase the analysis level of the source sentence taking into account its syntactic or semantic characteristics. Also, we think to parallelize our decoder. Indeed we use a bio-inspired algorithm that demands large computing time and memory. In SMT the most expensive step is the performance evaluation of hypotheses. This calculation is totally independent of one to another hypothesis. Whence it seems profitable to parallelize our decoder to improve quality and reduce costs.

## REFERENCES

1. Brown, P., Della Pietra, V., Della Pietra, S. & Mercer, R. 1993. The mathematics of statistical machine translation: Parameter estimation. *Computational Linguistics*, 19, 263-311.

2. Koehn, P. 2004. Pharaoh: A beam search decoder for phrase-based statistical machine translation models. In the *6th Conference of the Association for Machine Translation in the Americas* (pp. 115-124), *AMTA '04*, Washington, DC, USA.
3. Koehn, P., Hoang, H., Birch, A., Callison-burch, C., Federico, M., Bertoldi, N., Cowan, B., Shen, W., Moran, C., Zens, R., Dyer, C., Bojar, O., Constantin, A. & Herbst, E. Moses. 2007. Open source toolkit for statistical machine translation. In proceedings of the *45th Annual Meeting of the ACL on Interactive Poster and Demonstration Sessions, ACL '07* (pp. 177-180), Association for Computational Linguistics Prague, Czech Republic.
4. Crego, J. & Marino, J. 2007. Extending MARIE: an N-gram-based SMT decoder. In *Meeting of the Association for Computational Linguistics*. *ACL '07* (pp. 213-216), Prague, Czech Republic.
5. Niessen, S., Vogel, S., Ney, H. & Tillmann, C. 1998. A DP based search algorithm for statistical machine translation. In *Meeting of the Association for Computational Linguistics* (pp. 960-967), *ACL '98*, Montreal, Quebec, Canada.
6. Germann, U., Jahr, M., Knight, K., Marcu, D. & Yamada, K. 2001. Fast decoding and optimal decoding for machine translation. In *Meeting of the Association for Computational Linguistics* (pp. 228-235), *ACL '01*, Toulouse, France.
7. Ebadat, A., Smaili, K. & Langlois, D. 2009. Using Genetic Algorithm for the Decoder. Master's thesis, Saarland University, Computational Linguistics & Phonetics.
8. Watkins, A., Timmis, J. & Boggess, L. 2004. Artificial immune recognition system (airs): An immune inspired supervised learning algorithm. *Journal of Genetic Programming and Evolvable Machines* (pp. 291-317).
9. Abdelhadi, A., Mouss, H. & Kadri, O. 2010. Algorithmes du syst↕ eme immunitaire artificiel pour la surveillance industrielle. In *International Conference on Industrial Engineering and Manufacturing ICIEM'10*, Batna, Algerie.
10. Cettolo, M., Girardi, C. & Federico, M. 2012. Wit3: Web inventory of transcribed and translated talks. In proceedings of the *16th Conference of the European Association for Machine Translation (EAMT)* (pp. 261-268), Trento, Italy.
11. Och, F. & Ney, H. 2003. A systematic comparison of various statistical alignment models. *Computational Linguistics*, 29, 19-51.
12. Stolcke, A. 2002. Srilm-an extensible language modeling toolkit. In: proceedings of the *7th International Conference on Spoken*

*Language Processing* (pp. 901-904), *ICSLP '02*, Denver, Colorado, US.
13. Papineni, K., Roukos, S., Ward, T., j. & Zhu, W. 2002 Bleu: A method for automatic evaluation of machine translation. In *Meeting of the Association for Computational Linguistics* (pp. 311-318), *ACL '02*, Philadelphia, PA, USA.
14. Snover, M., Dorr, B., Schwartz, R., Micciulla, L. & Makhoul, J. 2006. A study of translation edit rate with targeted human annotation. In proceedings of *Association for Machine Translation in the Americas* (pp. 223-231).

**MANEL AMMAR**
INSTITUT SUPÉRIEUR D'INFORMATIQUE
ET DE MULTIMÉDIA DE SFAX,
SFAX UNIVERSITY,
SAKIET EZZIT, SFAX, TUNISIA.
E-MAIL: <MANEL1401@GMAIL.COM>

**SALMA JAMOUSSI**
INSTITUT SUPÉRIEUR D'INFORMATIQUE
ET DE MULTIMÉDIA DE SFAX,
SFAX UNIVERSITY,
SAKIET EZZIT, SFAX, TUNISIA
E-MAIL: <JAMOUSSI@GMAIL.COM>