

ALIGNING SENTENCES IN BILINGUAL TEXTS FRENCH - ENGLISH AND FRENCH - ARABIC

تقابل جمل لنصين مختلفي اللغة
فرنسي-انكليزي وفرنسي-عربي

Fathi DEBILI, Elyès SAMMOUDA
CNRS-idl

Conseil d'Etat, Palais Royal
75001 Paris, France

فتححي الدبيلي - إلياس صموده
المركز القومي الفرنسي للبحث العلمي

ABSTRACT

In this paper, we will tackle the problem raised by the automatic alignment of sentences belonging to bilingual text pairs. The method that we advocate here is inspired by what a person with a fair knowledge of the other language would do intuitively. It is based on the matching of the elements which are similar in both sentences. However, to match these elements correctly, we first have to match the sentences that contain them. There seems to be a vicious circle here. We will show how to break it. On the one hand, we will describe the hypotheses we made, and, on the other hand, the algorithms which ensued. The experiments are carried out with French-English and French-Arabic text pairs.

We will show that matching sentences and, later, expressions, amounts to raising a new problem in the machine translation field, i. e. the problem of *recognition* instead of that of translation, strictly speaking.

REMERCIEMENTS : Le travail présenté a bénéficié de l'aide de nombreuses personnes. Nous les remercions toutes, en particulier E. Souissi et A. Zribi pour leurs contributions ; E. Mackaay, I. Naddeo-Souriau, J.-L. Lemoigne et la revue *Pour la Science* pour la gentillesse avec laquelle ils ont accepté de nous donner sur disquettes des textes ou fragments de textes monolingues ou bilingues ; J. Kouloughli et J.-B. Berthelin pour les discussions et critiques qu'ils sont toujours prêts à faire.

Cette recherche a été en partie financée par le Réseau des Industries de la Langue (contrat ACCT n° 338/SG/C5) et en partie par le MRT (décision d'aide n° 90.K.6434).

ملخص

سنبحث في هذه المقالة مشكلة التقابل اللّغوي لجمل تنتمي إلى نصّين مختلفي اللّغة. الطريقة التي سنعمدها هنا مقتبسة مما يقوم به الفرد الذي يعرف لغة أخرى معرفة متوسطة ويرغب في إجراء هذا التقابل. هذه الطريقة تتركز على تقابل الكلمات المتماثلة في الجملتين المتناظرتين. ومن أجل تحقيق تماثل الكلمات هذا يجب أولًا مقابلة الجمل التي تحتويها. قد تبدو هذه المسألة حلقة مفرغة إذ أن تماثل الكلمات يتطلب تماثل الجمل وتقابل الجمل يتطلب تماثل الكلمات. نبيّن في هذه المقالة كيفية الخروج من هذه الحلقة المفرغة ونشرح الفرضيات التي اعتمدناها وكذلك الخوارزميات الموافقة. التجارب التي تم إجراؤها اعتمدت على أزواج من النصوص: فرنسي - انكليزي وفرنسي - عربي.

سنبيّن أيضًا أن مقابلة الجمل وفي المرحلة التّالية مقابلة العبارات تطرح مشكلة جديدة في مجال الترجمة اللّغوية وهي مشكلة التسرف لمد مشكلة الترجمة ذاتها.