LANGUAGE FORUM
VOL. 34, NO. 1, JAN-JUN 2008

A Classification Approach to Automatic Evaluation of Machine Translation Based on Word Alignment

KATSUNORI KOTANI
Kansai Gaidai University,Osaka, Japan

TAKEHIKO YOSHIMI
Ryukoku University, Shiga, Japan

HITOSHI ISAHARA
National Institute of Information andComm. Tech, Kyoto, Japan

TAKESHI KUTSUMI
ICHIKO SATA
Sharp Corporation, Nara, Japan

ABSTRACT

Constructing a classifier that distinguishes machine translations from human translations is a promising approach to automatic evaluation of machine-translated sentences. Using this approach, we constructed a classifier based on word alignment distributions between source sentences and human/machine translations, using Support Vector Machines as machine learning algorithms. We found that word alignment distributions succeeded both in achieving a classification accuracy as high as 99.4% and in identifying the qualitative characteristics of machine translations, which greatly helps improve the quality of machine translations.

1. INTRODUCTION

Previous research proposed a classification approach to machine translation evaluation in which a machine translation system can be evaluated based on the extent to which machine-generated translations (MTs) are similar to human-generated translations (HTs) (Corston-Oliver, Gamon & Brockett 2001; Gamon, Aue & Smets 2005; Kulesza