entity: an object or a set of objects in one of the semantic categories of interest
mention: a reference to an entity (typically, a noun phrase)
event trigger: the main word which most clearly expresses an event occurrence
event arguments: the mentions that are involved in an event (participants)
event mention: a phrase or sentence within which an event is described, including trigger and arguments

The 2005 ACE evaluation had 8 types of events, with 33 subtypes; for the purpose of this paper, we will treat these simply as 33 distinct event types. For example, for a sentence:

Barry Diller on Wednesday quit as chief of Vivendi Universal Entertainment.

the event extractor should detect a "Personnel_End-Position" event mention, with the trigger word, the position, the person who quit the position, the organization, and the time during which the event happened:

Trigger		*Quit*
Arguments	Role = Person	*Barry Diller*
	Role = Organization	*Vivendi Universal Entertainment*
	Role = Position	*Chief*
	Role = Time-within	*Wednesday*

Table 1. Event Extraction Example

We define the following standards to determine the *correctness* of an event mention:

- *A trigger is correctly labeled* if its event type and offsets match a reference trigger.
- *An argument is correctly identified* if its event type and offsets match any of the reference argument mentions.
- *An argument is correctly identified and classified* if its event type, offsets, and role match any of the reference argument mentions.

2.2 A Baseline Within-Sentence Event Tagger

We use a state-of-the-art English IE system as our baseline (Grishman et al., 2005). This system extracts events independently for each sentence. Its training and test procedures are as follows.

The system combines pattern matching with statistical models. For every event mention in the ACE training corpus, patterns are constructed based on the sequences of constituent heads separating the trigger and arguments. In addition, a set of Maximum Entropy based classifiers are trained:

- Trigger Labeling: to distinguish event mentions from non-event-mentions, to classify event mentions by type;
- Argument Classifier: to distinguish arguments from non-arguments;
- Role Classifier: to classify arguments by argument role.
- Reportable-Event Classifier: Given a trigger, an event type, and a set of arguments, to determine whether there is a reportable event mention.

In the test procedure, each document is scanned for instances of triggers from the training corpus. When an instance is found, the system tries to match the environment of the trigger against the set of patterns associated with that trigger. This pattern-matching process, if successful, will assign some of the mentions in the sentence as arguments of a potential event mention. The argument classifier is applied to the remaining mentions in the sentence; for any argument passing that classifier, the role classifier is used to assign a role to it. Finally, once all arguments have been assigned, the reportable-event classifier is applied to the potential event mention; if the result is successful, this event mention is reported.

3 Motivations

In this section we shall present our motivations based on error analysis for the baseline event tagger.

3.1 One Trigger Sense Per Cluster

Across a heterogeneous document corpus, a particular verb can sometimes be trigger and sometimes not, and can represent different event types. However, for a collection of topically-related documents, the distribution may be much more convergent. We investigate this hypothesis by automatically obtaining 25 related documents for each test text. The statistics of some trigger examples are presented in table 2.