1. Introduction

The Quran is a Muslim holy book that contains the word of God revealed to the Prophet Muhammad PBUH by the angel Gabriel to be read, understood, and practiced as a guide or a way of life for mankind[1]. The Quran consists of 30 juz, 114 surahs, 6236 verses and is written in Arabic and various other languages in the world[2]. The purpose of the translations is to make it easier for people to know the meaning of what is read and to better understand the meaning of the Quran.

In understanding the Quran by reading its translations, Muslims often find repetitive texts, both on the same surah or different surahs. It is "an indisputable fact"[3] that the Quran contains verses with repeated text. For example, *Maha Pemurah lagi Maha Penyayang*, is repeated 2 times in Surah al-Fatiha, that is in verse 1 and verse 3. By collecting verses that have similar texts, these verses can be analyzed for differences and emphasis over other verses[3]. As stated by Mahmoud Ayoub[4], "There are ... verses which are liable to more than one interpretation and so closely resemble one another in idiom and expression that they could lead those who are not firm in faith and knowledge to confusion and error." This example shows that collecting verses with similar texts is an important step for understanding the Quran correctly and accurately.

The current problem for searching longest repeated text in the Quran is it requires manual effort, by tracing the Quran verse by verse from the beginning to the end which takes a long time. This is a longest common substring (LCS) problem. Our solution for the system being made uses Ukkonen's algorithm[5]. This algorithm is used to build the longest repeated database of text sets in the Quran by comparing each verse with other verses which produces data containing the set of longest repetitive texts that are exact matches. This database is expected to be useful for further research to find the basic meaning between verses and obtain correct and accurate understanding of the Quran[3].