

Group 6

MovieProphecy

Software Design Document

Abhishek Chakraborty
Zachary Chapman
Yogesh Bodhe
Aravinda Kumar Yempada

Department of Computer Science
University of North Carolina at Charlotte

TABLE OF CONTENTS

1. INTRODUCTION	3
1.1 Purpose	3
1.2 Scope	3
1.3 Overview	3
1.4 Reference Material	3
1.5 Definitions and Acronyms	3
2. SYSTEM OVERVIEW	4
3. SYSTEM ARCHITECTURE	5
3.1 Architectural Design	5
3.2 Decomposition Description	6
3.3 Design Rationale	7
4. HUMAN INTERFACE DESIGN	8
4.1 Overview of User Interface	8

1. INTRODUCTION

1.1 Purpose

This report will define the design and technology decisions of our web based application called MovieProphecy.

1.2 Scope

MovieProphecy is an application for an average user to view the success or failure of a movie at the box office. It allows a registered movie executive to scrutinize the performance of a movie by displaying relevant data, charts and graphs.

1.3 Overview

This software design document defines and describes the use of each view, the architectural constraints of the system, the functional requirements with a significant impact on the architecture along with the layers and subsystems of the application.

1.4 Reference Material

The project of developing a movie prediction application called as MovieProphecy was inspired by a bunch of technical papers and proceedings on sentimental analysis of tweets while the most significant one being, '*Prediction of Movie Success using Sentiment Analysis of Tweets*'. (<http://www.jscse.com/papers/vol3.no3/vol3.no3.46.pdf>)

1.5 Definitions and Acronyms

API	Application Programming Interface – a set of routines, protocols and tools for building software applications.
IMDB	Internet Movie Database – a website that keeps records of movies, possible candidate for us to obtain lists of upcoming movies.
Java	Programming language used for the majority of our back end systems.
LingPipe	Java library for classifying text in numerous ways, used in our project for sentiment analysis.
SQL	Structured Query Language – special-purpose programming language for managing data held in a relational database management system.
Twitter	Social media site where users may publish 140 character posts for the world to see.
Twitter4J	Twitter API library for use with the Java programming language.
Tweet	A single post on Twitter.

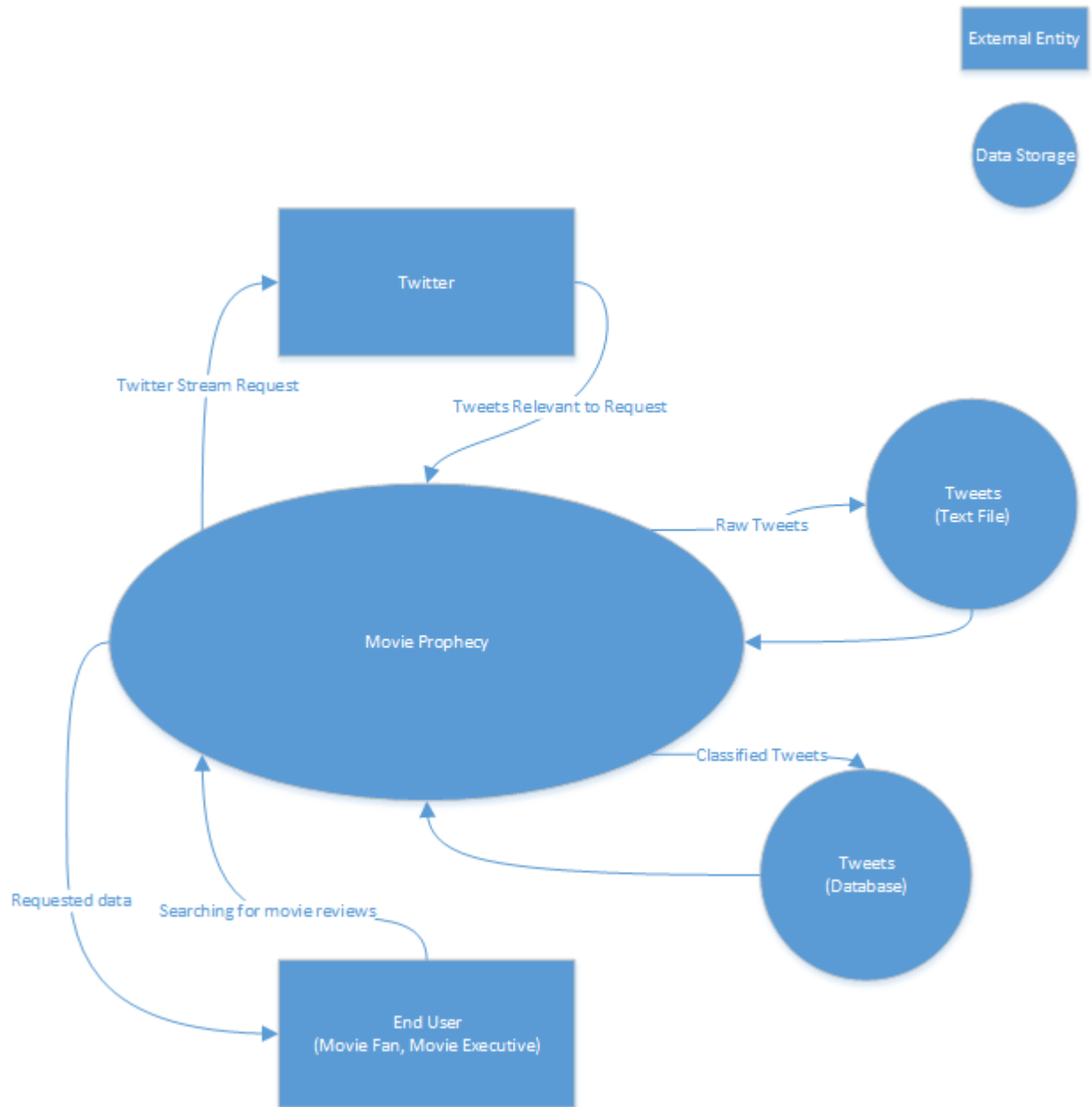
OAuth	Open Authentication – three-way handshake protocol used by the Twitter API to authenticate a user on Twitter and authorize them to access data.
OS	Operating System – software layer between the hardware and user applications.

2. SYSTEM OVERVIEW

We crawl the Twitter database to gather ‘movie tweets’ and are in particular, interested in the opinions that users have to share after watching a movie. Subsequently, we perform a sentiment analysis of ‘tweets’, the aim of which is to determine the attitude behind a user submitted written source. In a nutshell, the user is provided with a ‘crowdsourced’ review of a particular movie that the user maybe interested in.

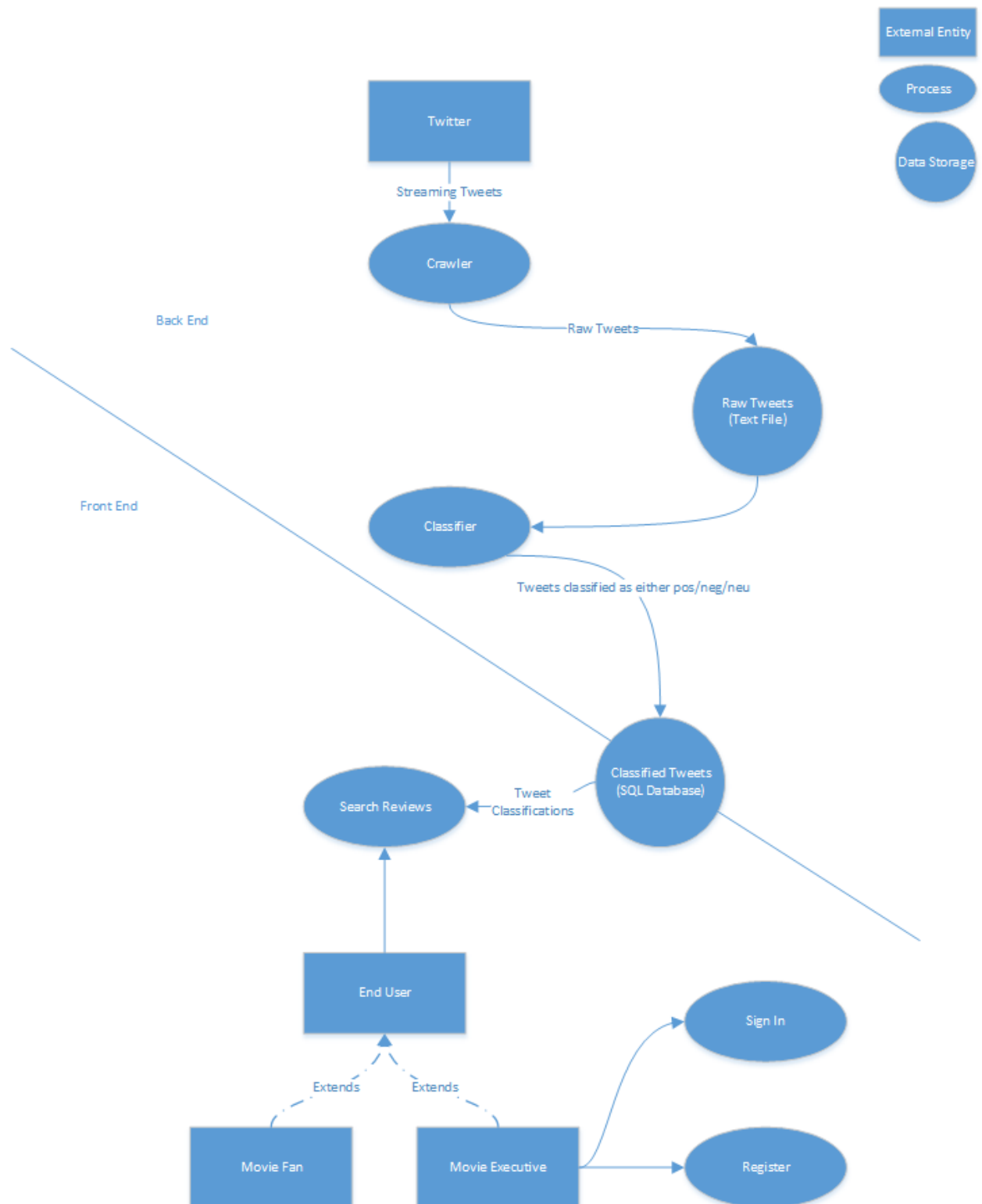
3. SYSTEM ARCHITECTURE

3.1 Architectural Design



The diagram depicts the various architectural modules used in MovieProphecy. The subsystems work as an integrated unit to gather movie centric tweets from the Twitter database and provide a sentiment analysis to end users who wish to gauge public reaction to a particular movie. The application uses a client-server architecture pattern where the majority of calculations are performed by the host server.

3.2 Decomposition Description



- Crawler – We crawl the twitter database using the stream and search API's for gathering tweets.
- Raw Tweets – The tweets gathered are then written onto a plain tab delimited text file.
- Classifier – We use Lingpipe 4.0 to sentimentally analyze the information gathered by our crawler.
- SQL Database – The processed and analyzed information is then stored into our SQL server.
- Frontend/ UI – A user searches for a movie through the frontend. Subsequently, relevant information about a movie and its reviews are pulled up and displayed on the user interface.

3.3 Design Rationale

We use the Twitter4J API for Java to crawl Twitter's streaming and searchable data. Java is a widely used and well documented programming language that many of us on the team have experience working with. Twitter's API provides us with the raw text of a Tweet as well as user, geological, and temporal metadata which is then stored in a tab delimited plain text file. Because some Tweets have missing attributes, using a SQL database here could potentially cause issues. Plain text is also easy to read and write using Java and does not have any storage complexities.

Raw Tweets are then sent to our classifier, which implements LingPipe 4.0's basic polarity classifier. LingPipe is written in Java, meaning it is easier to integrate into our system. LingPipe is also well documented and has been referenced in scholarly articles as a highly accurate classifier for full text movie reviews. Lastly, it is free to use for our current purposes. In the future we may need to either develop our own classifiers or pay LingPipe if we hope to commercialize our product.

Once analyzed, Tweets are written into a SQL Server 2012 database to be accessed by the front end system. Students have free access to SQL Server through Microsoft's DreamSpark initiative. Once again, we chose to use a well-documented relational database that acted as our comprehensive storage of polarized Tweets. Since our information is well structured, we went with SQL over NoSQL.

Our front end is created with ASP.NET. As with everything else, we went with comprehensive documentation. It is very user friendly due to drag and drop functionalities for interface design. It is also another tool that our team has hands on experience with from past projects.

4. HUMAN INTERFACE DESIGN

4.1 Overview of User Interface

This section describes the interactions of the user with our system in a step wise manner.

a. Home screen –

[Home](#) [About](#) [Register](#) [Log in](#)

Movie Prophecy

ABCDEFGHIJKLMNOPQRSTUVWXYZ

Upcoming

The Captive

Exodus: Gods and Kings

Inherent Vice

Top Five

The Color of Time

Recent

Penguins Of Madagascar

Big Hero 6

The Hunger Games: Mockingjay - Part 1

Interstellar

b. Alphabetical List (user selects 'A') –

Upcoming

The Captive

Exodus: Gods and Kings

Inherent Vice

Top Five

The Color of Time

Recent

Penguins Of Madagascar

Big Hero 6

The Hunger Games: Mockingjay - Part 1

Interstellar

Movie Prophecy

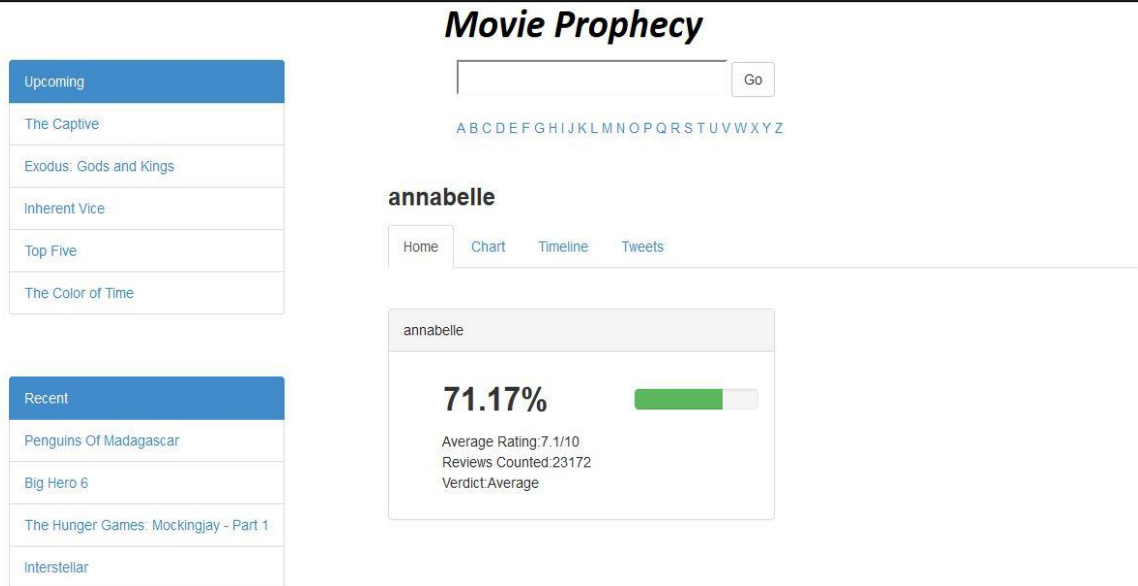
ABCDEFGHIJKLMNOPQRSTUVWXYZ

Movie Names Starting with A

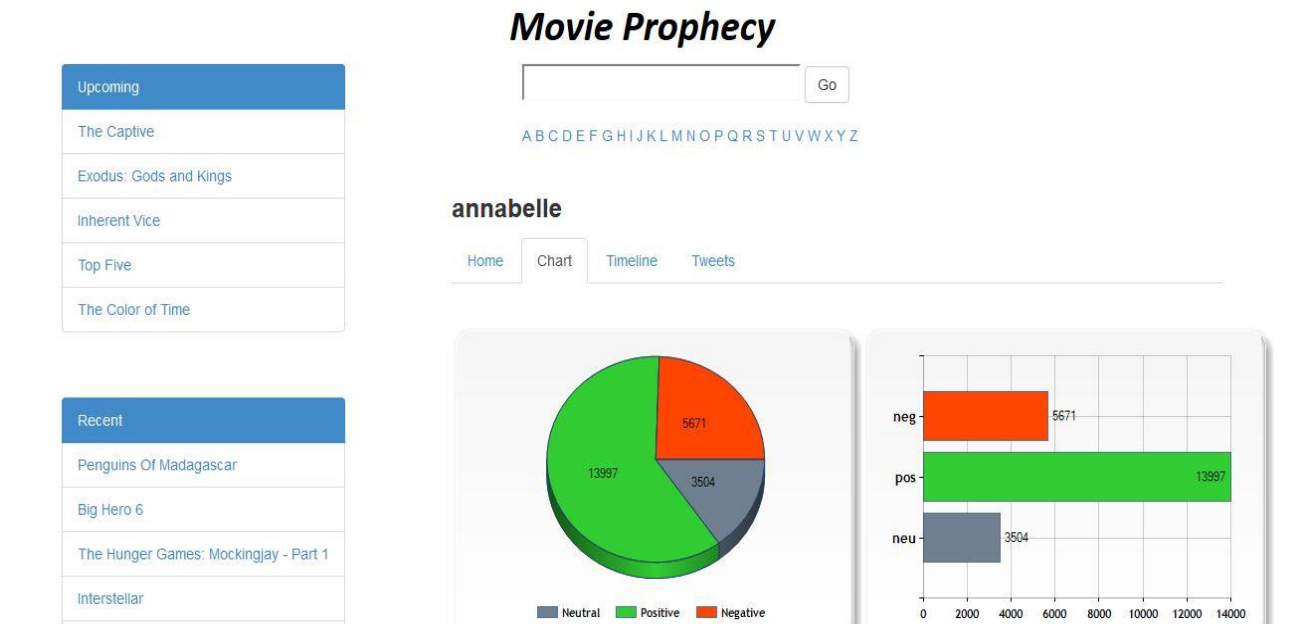
annabelle

abcs of death 2

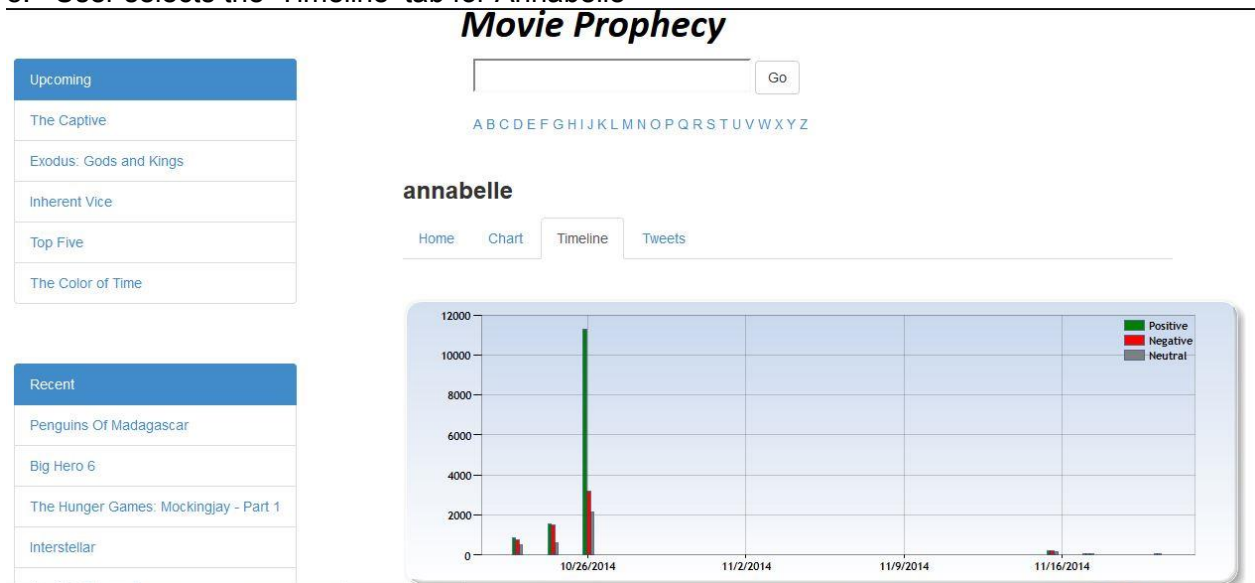
c. User selects 'Annabelle' –



d. User selects the 'Chart' tab for Annabelle –



e. User selects the 'Timeline' tab for Annabelle-



f. User selects the 'Tweets' tab for Annabelle-

Movie Prophecy

Upcoming

- The Captive
- Exodus: Gods and Kings
- Inherent Vice
- Top Five
- The Color of Time

Recent

- Penguins Of Madagascar
- Big Hero 6
- The Hunger Games: Mockingjay - Part 1
- Interstellar

annabelle

Home Chart Timeline Tweets

tweet_text	polarity
"@guiltarpleasure: Am I the only one who havent watch Annabelle yet ?" NO ?	neg
@nonnaechavez @angienynnb Movie marathon tah annabelle now playing but me not watching hahsha	neg
@jack_apperley whats Annabelle doing?	neg
@CourtsLeighn oh yeah I went to see annabelle last night dont go it shit scary man I had to sleep at byronys ???	neg
Annabelle again	neg
@swethhh11 annabelle made them do it	neg
I cannot believe Annabelle is still released... Urgh Im dead	neg