Offline Coding Task – Sponsored Search Auctions

1. Preliminaries and rules

Today's task is given by as large requirement text. *By purpose, it is very large and almost impossible to implement to the fullest extend within 3 hours*. It is up to you to define the scope! Please focus on the most important functionality first.

Please read the following introduction and requirements text first, so you know all requirements and design your software up-front to prevent major changes during the coding process. It is up to you to decide on the scope of the task - if you think you cannot implement all requirements, please focus on the important ones and decide which will not be implemented!

Proceed in an incremental, way to solve the task. Ensure that at any point in time, your solution can be executed and add small features one by one. In this way, the time limit will not hinder you to deliver at least a partial solution – which is very acceptable.

The maximum allowed time for completing this task is *3 hours*, which includes reading this description.

2. Definition of Done

The definition of done for any feature is:

- Unit- and integration tests are given that demonstrate the correctness of your code
- The code adheres to high quality standards (as few lines of code as possible), best practices and a clear structure. The code is easy to read without requiring comments and uses object oriented features such as inheritance and design patterns to make the code extensible and easy to maintain
- The resulting software is functional and can be tested by us

3. Introduction

Sponsored search is a type of web search engine advertising. Typically merchants try to place their web site offerings in a special region above or below the regular, algorithmic web search result list. The positions in those special, sponsored result region are more often than not sold in auctions, where merchants place bids for search terms. The highest bid wins and is displayed as the top most sponsored search result.

For example, when a user queries a web search engine for "stock trading", advertisers such as online banks may bid to have their listings featured alongside the regular, algorithmic search result listings. The advertisements appear in a separate section of the page designated as "sponsored" of the algorithmic results. The sponsored search results are displayed in a format similar to algorithmic results: as a list of items each containing a title, a text description, and a hyperlink to a corresponding web page. We call each position in the list a slot. Generally, advertisements that appear in a higher ranked slot (higher on the page) garner more attention and more clicks from users. Thus, all else being equal, merchants generally prefer higher ranked slots to lower ranked slots.

Advertisers bid for placement on the page in an auction-style format where the higher their bid's value, the more likely their listing will appear above other advertisements on

the result page. By convention, sponsored search advertisers generally pay per click, meaning that they pay only when a user clicks on their offering, and do not pay if their advertisement is displayed but not clicked. Google is the most prominent web search engine that employs this business model.

The sponsored search industry typically runs separate auctions for different search queries: for example, the queries “plasma television” and “investment advice” are associated with two distinct auctions. The entity being sold in each auction is the right to appear alongside the algorithmic search results of that search query. As mentioned, bids are expressed as a maximum willingness to pay per click.

For example, a fifty-cent bid by CHECK24 for “accident insurance” means CHECK24 is willing to pay up to fifty cents every time a user clicks on their advertisement. Advertisers may also set daily or monthly budget caps. In practice, hundreds of thousands of advertisers compete for positions alongside several millions of search queries every day. Generally the auctions are continuous and dynamic, meaning that advertisers can change their bids at any time, and a new auction clears every time a user enters a search query.

In this way, advertisers can adapt to changing environments, for instance by boosting their bids for the query “flowers” during the week before Valentine’s Day. The auctioneer, in this case the search engine, evaluates the bids and allocates slots to advertisers. Notice that, although bids are expressed as payments per click, the search engine cannot directly allocate clicks, but rather allocates impressions, or placements on the screen. Clicks relate only stochastically to impressions.

Search engines are an information gateway to many search and decision making tasks. As a result, entire niche industries exist touting services to boost a web page’s ranking on the popular search engines, in part by reverse engineering the search engines’ information retrieval algorithms. Research has shown that good placement on a search page leads to high traffic, and eventually an increased financial payoff. Paying for sponsored slots is an alternative means of obtaining prominent positioning. Sponsored search works because users often tolerate or even welcome targeted advertisements directly related to what they are actively searching for.

Typically, in sponsored search mechanisms, the advertisers specify a list of pairs of keywords and bids as well as a total maximum daily or weekly budget. Then, every time a user searches for a keyword, an auction takes place amongst the set of interested advertisers who have not exhausted their budgets.

Focusing on a single auction, let n be the number of bidders and $m < n$ the number of slots. The search engine estimates α_{ij} , the probability that a user will click on the i-th slot when it is occupied by bidder j. The quantity α_{ij} is called a click through rate (CTR). The CTR can be estimated from historic user behaviour. It is usually presumed for all j that $\alpha_{ij} >= \alpha_{i+1,j}$, for $i = 1, \ldots, m - 1$. In practice the CTR is estimated by the auctioneer and is not reported to the bidder. It is not clear that CTR’s can be estimated to a reasonable degree of accuracy, particularly by the bidder.

The search engine also assigns a weight w_j to each advertiser j. The weight can be thought of as a relevance or quality metric. If agent j bids b_j, his corresponding score is s_j

= $w_j b_j$. The search engine allocates slots in decreasing order of scores, so that the agent with highest score is ranked first, and so on. We assume throughout that agents are numbered so that agent j obtains slot j. An agent pays per click the lowest bid necessary to retain his position, so that the agent in slot j pays s_{j+1}/w_j.

This weighted bid ranking mechanism includes the two most prominent keyword auction designs that have been used in practice: the "rank by bid" mechanism ($w_j = 1$) and the "rank by revenue" mechanism ($w_j = \alpha_{1j}$). Both variants are sometimes called generalized second price auctions. In the generalized first price auction, agents are ranked by bid but each bidder who secures a slot pays their bid per click.

An auctioneer usually has the objective to maximize revenue. Under this objective, the auctioneer needs to determine which allocation of bidders to slots gives the maximum expected revenue, which is a straightforward task, if we do not include any other information.

4. Requirements and Task

Implement a simple, sponsored search auction system that implements all four described auction mechanisms - the two rank by bid and the two rank by revenue mechanism.

For a given period of time, say a day, or a week, the auction system executes auctions in a strictly sequential order. The execution of an auction is triggered once a users searches for a search term for which at least one bidder has registered at the auction system. A bidder holds a private set of search terms that are registered with the auction system once a bidder registers at the system. For each search term, a bidder has a fixed maximum bid price that only the bidder knows of (the maximum bid price may even change over time). A bidder also has a budget for each period of time that is also submitted to the auction system. The auction system makes sure, that the budget is not overrun, which may only occur, if a user clicks on the advertisement a bidder has been bidding for in the list of sponsored search results. Consequently, the budget is only charged if an advertisement has lead to a user's click. A bidder may only win a single slot in any auction.

In an auction, a fixed amount of bidders is participating, competing for the slots in the sponsored search results for a single user. For each bidder, the auctioneer determines the required data for w_j and α_{ij} before any auction is executed. The auctioneer also defines which auction mechanism shall be executed. w_j and α_{ij} directly influence on the results of an auction. The sponsored, ordered result list is presented to the user who decides on which list entry to follow. A user may also decide not to click on any sponsored results in the list.

Please provide unit- and integration tests to drive your software; we do not require any user interface! Focus on the domain model and the verification of the provided functionality.