1. What is the definition of an information retrieval system?

An information retrieval system (IRS) is a system that helps users find information from a collection of documents. IRSs typically use a variety of techniques, such as keyword searching, to help users find the information they are looking for.

2. Contrast the differences between IRS and DBMS?

An information retrieval system (IRS) is a system that helps users find information from a collection of documents. A database management system (DBMS) is a system that stores and manages data. IRSs and DBMSs are both computer systems that store and manage information, but they have different purposes. IRSs are designed to help users find information, while DBMSs are designed to store and manage data.

3. Illustrate the Inverted File structure and PAT Data structure?

An inverted file is a data structure that stores information about the terms that appear in a document. Each term is associated with a list of documents that contain the term. The PAT data structure is a data structure that stores information about the patterns that appear in a document. Each pattern is associated with a list of documents that contain the pattern.

4. Multimedia Database Search

Multimedia database search is a type of information retrieval that focuses on finding information from multimedia documents, such as images, videos, and audio files. Multimedia database search is more challenging than traditional information retrieval because multimedia documents are more complex and can contain a variety of different types of information.

5. Explain the terms i) Recall ii) Precision

Recall is the fraction of relevant documents that are retrieved by a search system. Precision is the fraction of retrieved documents that are relevant. Recall and precision are two important measures of the performance of an information retrieval system.

6. Explain the Total Information Retrieval System?

A total information retrieval system (TIRS) is a system that attempts to retrieve all relevant documents from a collection. TIRSs are typically used in situations where it is important to retrieve all relevant documents, even if this means retrieving a large number of irrelevant documents.

7. Explain Digital Libraries and Dataware House?

A digital library is a collection of digital objects, such as text, images, videos, and audio files. Digital libraries are typically organized and managed using information retrieval techniques. A data warehouse is a collection of data that is used to support decision-making. Data warehouses are typically organized and managed using database management techniques.

8. Explain the conditions of Porter Stemming Algorithm —_?

The Porter stemming algorithm is a stemming algorithm that is used to reduce words to their root form. The algorithm is based on a set of rules that are applied to words. The rules are designed to remove suffixes from words.

9. Explain the PAT Data structure? — Pg. No. [88-93]

The PAT data structure is a data structure that stores information about the patterns that appear in a document. Each pattern is associated with a list of documents that contain the pattern. The PAT data structure is used to improve the performance of information retrieval systems.

10. Two-level Bayesian networks with neat diagram pg. No. [61]

A Bayesian network is a graphical model that is used to represent the joint

probability distribution of a set of variables. Bayesian networks are used in a variety of applications, such as information retrieval, natural language processing, and machine learning.

11. Explain

(a) Canned Query pg. No. [43]

A canned query is a query that is pre-defined and stored in a database. Canned queries are used to improve the performance of information retrieval systems.

(b) Bigrams, Trigrams and Pentagrams [85]

A bigram is a sequence of two words. A trigram is a sequence of three words. A pentagram is a sequence of five words. Bigrams, trigrams, and pentagrams are used in information retrieval to improve the performance of search engines.

(c) N-Gram [85]

An n-gram is a sequence of n words. N-grams are used in information retrieval to improve the performance of search engines.

12. What are Classes of Automatic Indexing pg No. [106]

There are three main classes of automatic indexing: statistical indexing, rule-based indexing, and hybrid indexing. Statistical indexing uses statistical methods to create indexes. Rule-based indexing uses rules to create indexes. Hybrid indexing uses a combination of statistical methods and rules to create indexes.