•What we're looking for (~10 minutes of presentation)

**–You understood the paper**

–Discussion of related work and **background**

–Discussion of why should we care about the topic, and more importantly why your fellow classmates should care & –Relation of your paper to the lecture slides I gave on the topic

–Simple summarization and description of the algorithm and/or technology introduced in the paper

–What were the results/contributions/conclusions of the paper

–Your evaluation of Pros of the paper

–Your evaluation of Cons of the paper

•What we're NOT looking for

–Plagiarism

–Repetition

–Cutting/Pasting inappropriately out of the paper

–Regurgitation

–You to follow the EXACT set of bullets that we gave on the prior list

-You should be looking to be innovative – show the class and instructors that you really understood what was in the paper - Treat it like a conference presentation

Thanks,

整理

* Understood of the paper / about paper itself:
* intro:

This article was written to introduce Envision, a multimedia digital library of computer science literature and it mainly focuses on its ability to display and visualize the search results of the query entered by the user.

According to this paper and other related publications, I find that the most resources used in Envision database are **bibliographic records** from ACM Guide to Computing Literature, **review articles** from Computing Reviews, **full-texts** supplied by the typesetter of the Transactions, and **full-texts** keyed in of Communications. It also has videos and animations and that’s why it is called multimedia.

The purpose of the paper is to introduce Envision and the purpose of Envision project is to enhance the ability of computer science students to search the literature for solutions. Since at that time, authors think students in cs area dislike reading, have little exposure to technical reports or articles in the research literature, and so are not prepared for lifelong learning in their profession. So they tried to develop this project to help!

* Go through the paper:

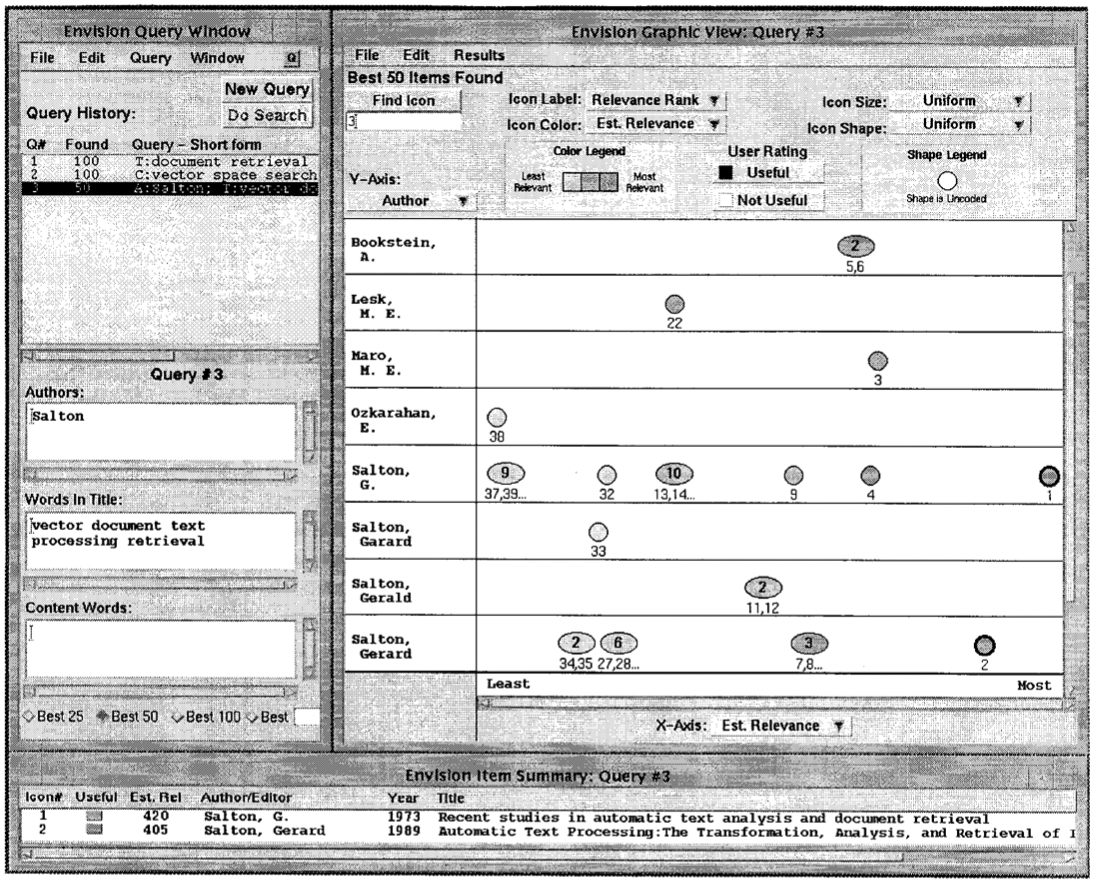
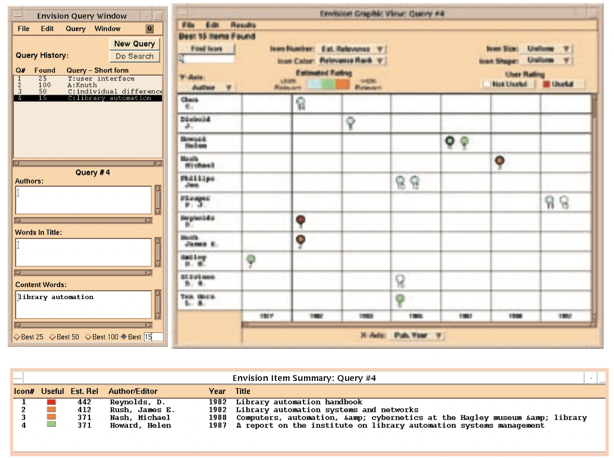
Let’s go through the paper briefly. The authors divided this paper into 5 parts. First, they introduced the basic information about Envision and compared it with several function-like software.

In second part titled “Development of Envision”, authors mainly described that they interviewed some professional people in target area, and according to what they obtained they decided to develop additional system of information visualization which is different from the initial plan of Envision.

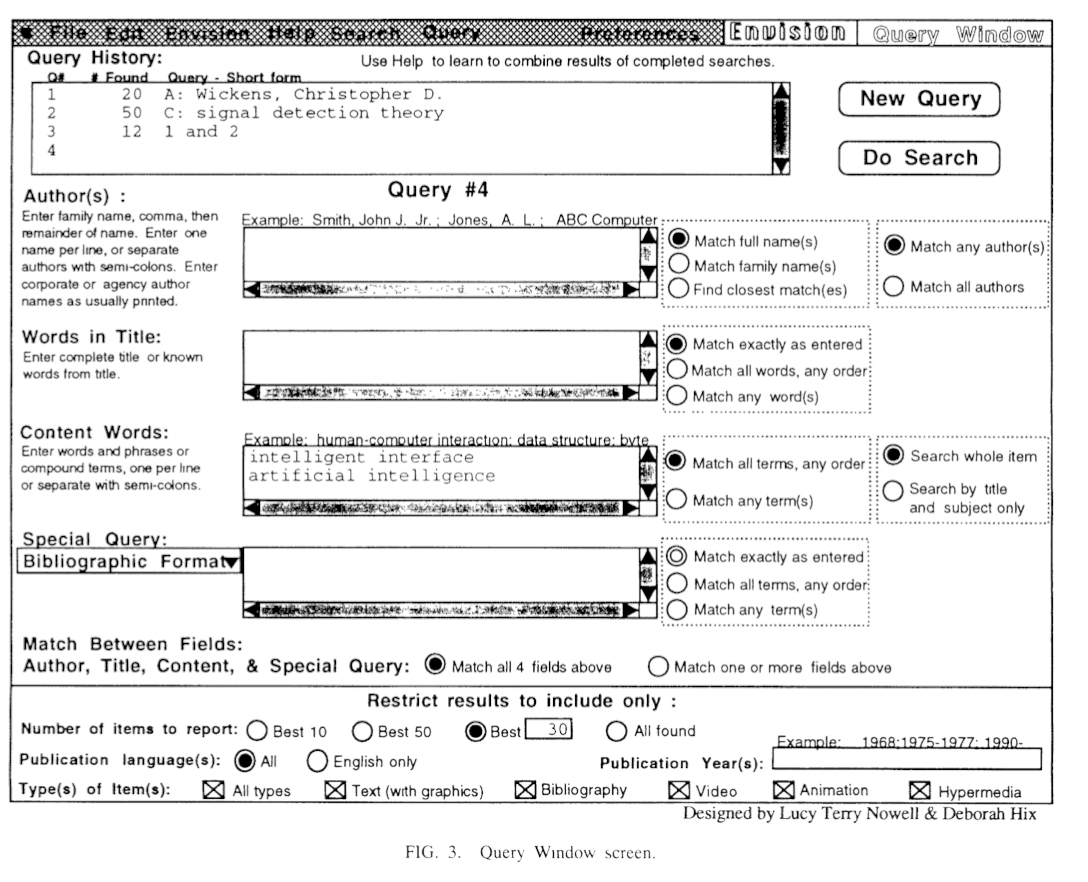
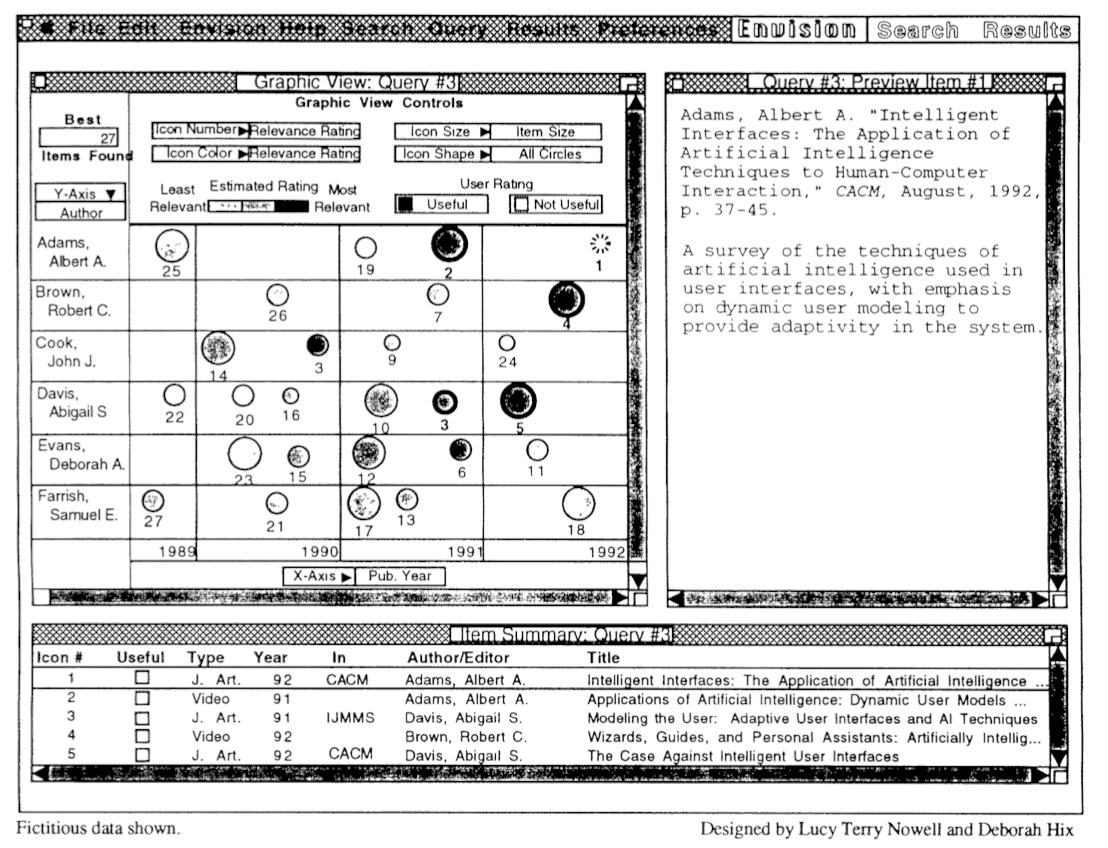
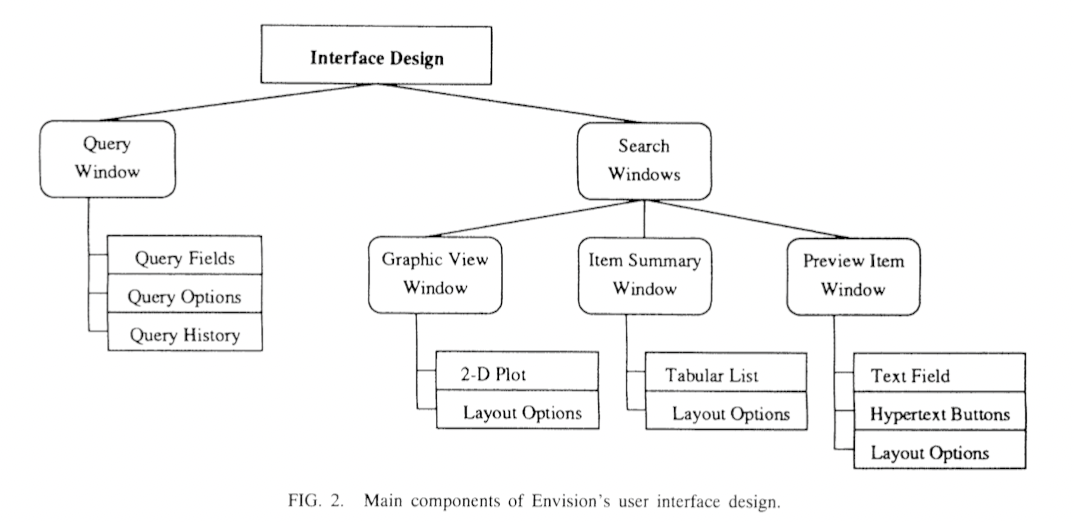
And then, the third part, which is the most important part of this paper. It shows the design for Visualization system of Envision, like what are the three windows of Envision and what do they mean, how about those semantics shown and how to make your own semantics. Besides, authors showed several example views of different queries and semantics, and they also explained some underlying information and analysis.

For the user interface, window on the left, called Query window, is used for user interaction. We, user can enter the conditions we want to apply on the retrieval results and limit the records returned. The one on the bottom, named Item Summary window, shows the summary information about data retrieved. And the most outstanding one, called Graphic View window, is used to display the visualization of the search results and also change the options of plot.

And about the visualization designs, the attributes of documents include author name, publication year, document type, estimate document relevance, index term and etc. Just take one as example, as the **picture** showing, by choosing author name on the y-axis, estimated relevance on the x-axis and relevance rank shown by the icon label. it identifies which author has produced how many works relevant to the specific topic. The icon color also can give information about the relevance, we can see the **picture at right,** documents in the top 35% of relevance values are assigned to be orange, the next 35% are green and the rest are blue.



With years of development, said by one publication in 2006, there was an additional Preview Item window placed with Item summary window and Graphic View window(see **picture xx**), while the Query window was assigned to a single window in a page with more details(see **picture xx**)



The fourth part describes the result of usability test among participants in the area of computer science and the last part summarizes the status of Envision at that time and authors’ expectation and brief solutions to some limitations.

* Simple summarization and description of the algorithm:

I think there are two major algorithms used in the Envision project, the first is Query-document similarity computation algorithm to quantify the estimated document relevance. Unfortunately, I didn’t find any information about the method authors used since nearly all referential documents of this paper are about information visualization, not document similarity. And the rest are about Envision itself and don’t mention what I want as well. I will be appreciated if professor or anyone can make a supplement.

And the second is Graphic view layout algorithm. Before we talk more, we should know that some attributes with discrete values, like author and publication year, can divide the window into several cells and each retrieved document is represented by a single icon. Since there are multiple choice of these discrete-value attributes, resulting in a document may belong to multiple cells and a cell can have numbers of icons. So this algorithm can decide how are the icons displayed within a cell and The principles of it are shown in picture. To meet these principles, the authors make an elliptical cluster icon to represent documents that are very close to each other and the number of documents is shown within the elliptical cluster while the label under cluster reveals the rank 文本

描述已自动生成of the top two relevant documents.

* results/contributions/conclusions/Pros: what this paper contributed

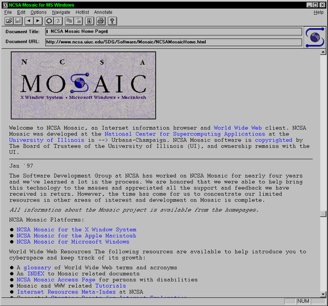
All in all, this paper shows a clear picture of Envision project for us, including its function, design details, limitations and possible solutions. I think it can be considered as a very successful project of the information visualization system in the early age. There are many latter publications have cited this paper as part of their inspiration for information display.

Nowell et al. [22] present a highly flexible way of organizing search results. Users can organize and explore result sets in the cells of a matrix view whose axes can be set to any aspect of the documents, e.g. author and year of publication. Additional metadata attributes can be included by mapping them to shape, pictograms, label, or color of the respective document icon in its matrix cell.

Nowell et. al [18] propose a more powerful scatterplot-based method that combines glyphs, a coloring scheme and icons in a customized search interface

* Cons of the paper:

But this paper was written in 1996 and the latest relative paper by Nowell, lead author of this paper, was published in 2006, so it is somewhat an old paper for us.

For example, the browser it used is Xmosaic, and its full name is Mosaic for the X Window System. Although it is the first commonly used web browser which can display images inline with text, UIUC and NCSA, the designers of it, decided to not keep it maintained and developed in 1997.

And as we talk in algorithm part, I think one of the drawbacks of this paper is that it didn’t explain the algorithms it used very clearly, especially how to compute the value of relevance, because it is a key attribute in Envision. And it is also related to knowledge we learned in class.

Besides, the limited number of participants joined in the evaluation of Envision also can be a shortcut. The author only invited 5 people to test this project in 1993 and didn’t make additional test with more people to join in during the 3 years before writing this paper. I think the number of testing people is not large enough to make the feedback from them significant.

* why should we care about the topic, and more importantly why your fellow classmates should care & Relation of your paper to the lecture slides I gave on the topic:

In fact, what this paper talking about is kind of relative to our study. In this class we are learning about data analysis and representation better, especially for document content. I think it may be inspirational for us to learn to obtain diverse results by changing different representing scales. However, I believe most of us must be familiar with this since this is a paper nearly 30 years ago and data analysis system has matured now.

text analysis

稿子:

This article was written to introduce Envision, a multimedia digital library of computer science literature and it mainly focuses on its ability to display and visualize the search results of the query entered by the user.

The purpose of this project is to enhance the ability of computer science students to search the literature for solutions. Since at that time, the authors think students in cs area dislike reading, have little exposure to technical reports or articles, and so are not prepared for lifelong learning in their profession. they decided to develop this project to help!

According to this paper and other related publications, I find that the data used are bibliographic records from ACM Guide to Computing Literature, review articles from Computing Reviews, full-texts supplied by the typesetter of the Transactions, and full-texts keyed in of Communications. It also has some videos and animations and that’s why it is called multimedia.

Let’s go through the paper briefly. The authors divided this paper into 5 parts. First, they introduced the basic information about Envision and compared it with several function-like software.

In second part titled “Development of Envision”, authors mainly described that they interviewed some professional people in target area, and according to what they obtained they decided to develop additional system of information visualization which was different from the initial plan.

And then, the third part, which is the most important part of this paper. It showed the design of Visualization system, like what are the three windows of Envision and what do they mean, how about those semantics shown and how to make your layout. Besides, authors showed several examples and explained some underlying information and analysis.

The fourth part described the result of usability test among participants in the area of computer science and the last part summarized the status of Envision at that time, authors’ expectations and brief solutions to some limitations.

Now we go back to the design part and see more details.

For the user interface, window on the left, called Query window, is used for user interaction. We can enter the conditions we want to apply on the retrieval results and limit the number of records returned. The one on the bottom, named Item Summary window, shows the summary information about data retrieved. And the most outstanding one, called Graphic View window, is used to display the visualization of search results, and also can change options of the plot.

And about the visualization designs, the attributes of documents include author name, publication year, document type, estimate document relevance, index term and etc.. All these attributes can be chosen in pop-up windows in layout option area and x and y axis. Just take one as example, as the picture showing, by choosing author name on the y-axis, estimated relevance on the x-axis and relevance rank shown by the icon label. It identifies which author has produced how many works relevant to the specific topic.

The icon color also can give information about the relevance, we can see this colorful one, there are three colors shown. Documents in the top 35% of relevance values are assigned to be orange, the next 35% are 表格

描述已自动生成green and the rest are blue. And these can also be found in Item Summary window.

With years of development, said by one publication in 2006, there was an additional Preview Item window placed with Item summary window and Graphic View window, while the Query window was assigned to a single window in a page with more details.

And as for the design algorithms, I think there are two major algorithms used in the Envision project, the first is Query-document similarity computation algorithm to quantify the estimated document relevance. Unfortunately, I didn’t find any information about the method authors used since nearly all referential documents of this paper are about information visualization, not document similarity. And the rest are about Envision itself and don’t mention what I want as well. I will be appreciated if professor or anyone can make a supplement.

And the second is Graphic view layout algorithm. Before we talk more, we should know that some attributes with discrete values, like author and publication year, can divide the window into several cells and each retrieved document is represented by a single icon. Since there are multiple choice of these discrete-value attributes, resulting in a document may belong to multiple cells and a cell can have numbers of icons. So this algorithm can decide how are the icons displayed within a cell and The principles of it are shown in picture. To meet these principles, the authors make an elliptical cluster icon to represent documents that are very close to each other, and the number of documents is shown within the elliptical cluster while the label under cluster reveals the rank of the top two relevant documents.

All in all, this paper shows a clear picture of Envision project for us, including its functions, design details, limitations and possible solutions. I think it can be considered as a very successful project of information visualization system in the early age. There are many latter publications have cited this paper as part of their inspiration for information display.

However, There are some drawbacks of this paper. It was written in 1996 and the latest relative paper by Nowell, the lead author of this paper, was published in 2006, so it is somewhat an outdated paper for us.

For example, the browser it used is Xmosaic, and its full name is Mosaic for the X Window System. Although it is the first commonly used web browser which can display images inline with text, UIUC and NCSA, the designers of it, decided to not keep it maintained and developed in 1997.

And as we talk in algorithm part, I think one of the drawbacks of this paper is that it didn’t explain the algorithms it used very clearly, especially how to compute the value of relevance, because it is a key attribute in Envision. And it is also related to knowledge we learned in class.

Besides, the limited number of participants joined in the evaluation of Envision also can be a shortcut. The author only invited 5 people to test this project in 1993 and didn’t make additional test with more people to join in during the 3 years before writing this paper. I think the number of testing people is too small to make the feedback significant.

In fact, what this paper talking about is kind of relative to our study. In this class we are learning about data analysis and representation better, especially for document content. I think it may be inspirational for us to learn to obtain diverse results by changing different representing scales. However, I believe most of us must be familiar with this since this is a paper nearly 30 years ago and data analysis system has matured now.