Exchange of mathematical information on the web: Present and Future

Manasi Athale
Rahul Athale
Research Institute for Symbolic Computation,
Johannes Kepler university, Linz, Austria
E-mail: manasi,athale@risc.uni-linz.ac.at

Abstract

The advancement in the technology has given us the new and most powerful communication medium, Internet. It has opened up many new avenues and also has posed different problems. These can be categorized into two broad categories, namely problems regarding storage and problems regarding communication. This can be explained in scientific terms as computer-supported build-up of mathematical knowledge and computer-supported retrieval of mathematical knowledge.

In this paper, we suggest that for storing data, we categorize the users according to their needs, or in other words, according to their usual usage of information. We explain our categorization based on the needs with specific examples and why we came up with these categories. To fulfill the needs of users from different categories, we either have to store the same information in many formats with minor changes, or add some extra information, using which we can use the same information for different users. We give an example to show our approach in this regard.

1 Introduction

The advancement in the technology has given us the new and most powerful communication medium, Internet. Internet has many advantages, for example, it is cheaper than talking on the phone, it is faster than any other communication media. Thus, it's inexpensiveness and easy availability has made it very popular. This has benefitted the scientific community but has also introduced common people to the vast knowledge from all fields of human endeavor. Today even layman can seek the answers of his questions about everyday weather in his home-town using the Internet. The High-school student can look for the help on the solutions of his homework.

There are two ways to seek an answer for a problem on the web. One way to get the answers from the web is to search for that answer on the web. The second choice is to pose the question to experts in the field using the websites such as, [3, 5] [1]. This communication has certainly helped the students (we are not looking into the ethical and moral issues originating from this communication), but it has posed new questions. The first option, to search for the answer, is related to storage while the next one, to ask for the answer from experts, is related to communication. This can be explained in scientific terms as computer-supported build-up of mathematical knowledge and computer-supported retrieval of mathematical knowledge. Of course both things are intricately intertwined. Let us ponder over this for a moment.

Suppose a student asks a question.

Question: Is 2345 divisible by 3?

He can search the knowledge base already available on one of the sites which were earlier mentioned, but it is very hard to find a readymade answer for such a specific problem. Now he asks this question to a group of experts, which can give rise to very many answers, let us list some representatives:

Answer 1: Please use your calculator.

(Gives a method which can be generalized but no mathematical content.)

Answer 2: $2+3+4+5=14$ as 3 does not divide $14, 3$ does not divide 2345.

(Gives rule but very hard for usual student to generalize.)

Answer 3: A number is divisible by 3 if and only if sum of its digits is divisible by 3. Now use this to answer your question.

(Good motivation but difficult for a student who is not good at Mathematics.)

Answer 4: Do not ask such silly questions.

(No comments!)

Now second student from the same class asks another question which is just a reformulation of the previous one.

Question: State divisibility test for 3.

This student still will not be able to get the answer from the responses to first question which are already stored, unless they are again parsed and

[1] One of the author of this paper devotes some of his free time to help users with Mathematical questions on this site.

categorized by another expert in the field who only devotes his time for this task. There are many questions which come to mind after reading this, like why we are looking at this particular problem, what is its use etc. Let us state the problems which will be addressed in this paper.

2 Problem description

"Every workman in the exercise of his art should be provided with proper Implements." Introduction to Roget's Thesaurus by P.M. Roget [6]. As Mathematics is the language of all sciences, this statement also applies to Mathematics. Roget, in his idea of thesaurus, tried to solve the problem of grouping the words from the vast set of English language words. His main aim was to provide authors and translators the words with similar shades at the same place. In this retrospect, our problem is similar, how to group the available mathematical information on a specific topic, for the easy access to help all the users at the same time. Here we would like to point out that *American Mathematical Society* has already categorized the subfields in Mathematics and some other people like Rusin [8][2] have put their own classification on the web. We are not suggesting classification of this kind.

In the introduction, we came across *two* questions, where for a mathematician, answer to the first will be just application of the second answer. This leads us to the first problem, *why we are considering students, or for that matter, laymen in our examples.*

Every year thousands of books and research articles are published on Mathematics. Many basic things are repeated in different contexts, for example, basic algebra, topology and analysis books, all have a chapter on set theory. When it comes to the electronic storage and access of the information, one should try to minimize the repetition of information as new information is growing exponentially and the storage space is finite. *How we can categorize the information, so that, it can cater to the different needs of laymen as well as scientists, high-school students as well as mathematicians and so on?* This gives rise to the next question *how to present the information using the categorical building blocks developed earlier.*

The present situation, as explained in the introduction, is not very promising when it comes to providing personalized advice for special needs. *What can and cannot be done in near future?* In other words, can we provide personalized mathematical information using given infrastructure with some minor improvements?

[2] This has a very good explanation which is accessible to laymen also.

3 Categorical approach

3.1 Why we are considering students, or for that matter, laymen?

In his recent paper [10], Kaltofen points out that "The students learning calculus with the help of a computer algebra system constitute the most numerous users of our software". Similarly when we consider exchange of mathematical information on the web, there will be many more users accessing basic information than the advanced one. This does not mean that the common people will not access the information on advanced topic. Lang [11] has already demonstrated that, if the subject is well presented, even *topology* can be taught to a person with no training in Mathematics. We are considering here the general trends. Therefore, we have to make efforts to present the information in nice format to our novice consumers. As these users will form the largest part of all the users of our knowledge base they will bear almost all the cost of maintenance of it. The mathematicians as the developers of this knowledge base will naturally make it in a way which is useful to all their community.

3.2 How we can categorize the users?

There are *two* issues involved, namely how to decide the categories according to the needs of the users and then how to categorize the information to fulfill them. Let us tackle them one after the other.

Category	**Appreciates**
layman	a real life example
school student	a real life example with mathematical statement
engineer	main result, emphasis on the application part
other scientist	main result, application, links to similar results
mathematician and/or graduate math student	rigorous proof, theoretical examples, similar results, links to research papers

Table 1: Broad categories of users

The table 1 shows a broad categorization of users and the usual category of information they seek when they search the Internet or the knowledge base. This is a very broad categorization but let us see why this is relevant. We have come up with these categories with our own observation and discussions with fellow users, say about thousand of them. Let us see which kind of users belong to each category.

Layman: Layman represents group of users who have no formal training in Mathematics. The people who studied Mathematics but never used it, except in everyday shopping, constitute the big part of this group. Therefore these people will appreciate a real life example explaining a mathematical concept, which may be, they want to teach to their five year old. A link to theoretical example might be helpful.

School student: School student represents students in the age of 10–16. They know few basic things but to make things clear to them real life example is the best. Most probably they will be learning the concept in their maths class, so they would like to see the mathematical formulation. A link to theoretical example might be helpful.

Engineer: Engineer represents users who apply Mathematics in their work. They want to know the mathematical result and it's application. A theoretical example well-worked out might be useful. If an algorithm exists then a link to it will be must.

Other scientist: Other scientist represents who are scientists but not mathematicians. They always use Mathematics so for them main result and application is must. As they are into research, link to similar results will be certainly appreciated.

Mathematician: Mathematician represents users who are working mathematicians or budding ones. They are the people who make most of the contributions to knowledge base. They want to know the rigorous proof of the result to search for the possible generalizations. Therefore theoretical examples and similar results are must. Links to research papers is a must for researchers in that specific field.

The categories are not exclusive and one has to choose the categories according to the information he is looking for. For example, though we know little bit of topology, it would be interesting to know how it is explained to layman, then we can choose layman and search for topology. An engineer might want to recall the proofs of results he is using then he can change to mathematician mode. We can also provide links to further information in each category according to further interests.

3.3 How we can categorize the information?

Now we have the categories of the users and we want to divide the information. This we will explain using an example which is, as simple as possible but as complicated as necessary,[2]. We will take the example based on the *two* questions in the introduction.

Divisibility test for 3 :

Real life example: Suppose we have 2345 apples and we want to distribute them between 3 vendors equally, is it possible? Let us take out 3 apples from 2345, we remain with 2342 apples. If we continue in this way, we will be left with 2 apples, which cannot be distributed equally amongst 3 people. Therefore we cannot distribute 2345 apples equally between 3 vendors. Mathematically speaking, we were trying to find out whether 2345 is divisible by 3 or not. Subtracting repeatedly is same as dividing. This you can check using your calculator but you can do it in your head using a simple test.
Mathematical rule or test: *[Divisibility test for 3]* A number is divisible by 3 if and only if sum of digits of the number is divisible by 3.
Theoretical examples:

- Check if 2345 is divisible by 3.
  Sum of digits of 2345 is $2 + 3 + 4 + 5 = 14$. Applying the rule we get, as 3 does not divide 14, 3 does not divide 2345.

- Check if 3 divides 789.
  Sum of digits of 789 is $7 + 8 + 9 = 24$. Applying the rule we get, as 3 divides 24, 3 divides 789.

- Check if 123456789999 is divisible by 3.
  Sum of digits of 123456789999 is 72. As 72 is slightly bigger let us check whether 3 divides it or not using our rule. Applying the rule we get, as 3 divides $7 + 2 = 9$, 3 divides 72. Applying the rule again, as 3 divides 72, 3 divides 123456789999.

Take few numbers and carry out the above procedure.
Proof: Let us suppose that our number $n = a_l * 10^l + \ldots + a_1 * 10 + a_0$. Now we have to prove that 3 divides n if and only if 3 divides the sum $m = a_l + \ldots + a_1 + a_0$. Let us suppose that 3 divides m to prove that 3 divides n.

$$
\begin{aligned}
n &= a_l * 10^l + \ldots + a_1 * 10 + a_0 \\
&= (a_l * (10^l - 1) + \ldots + a_1 * (10 - 1)) + (a_l + \ldots + a_1 + a_0) \\
&= 9 * \underbrace{(a_l * (10^{l-1} + 10^{l-2} + \ldots + 10 + 1) + \ldots + a_1)}_{p} + m \\
&= 9p + m
\end{aligned}
$$

As 3 divides 9 and 3 divides m, 3 divides $n = 9p + m$. Similarly we can prove in the other direction.
Application: We can use the test to find out new primes, that is, in primality testing. If a number is divisible by 3 then it is not prime.
Generalizations:

- The same proof can be generalized in case of 9. A number is divisible by 9 if and only if sum of digits of the number is divisible by 9.
- How can we generalize divisibility test for 3, for numbers represented in bases other than 10.

Please visit webpages [7, 9] for good examples on divisibility tests.

Presentation of the information:
The above categorization saves space, as the mathematical rule is shown to all the users but is written only once. Many of the parts are used for more than one category of users. Usually on all the websites, we observe that either they lack proof or examples. After these categorizations, we will add different tags to these information modules. We will display the information as per the choices made by the user. Now if one tries the *two* questions from introduction, or similar such questions, he will get the required information.

3.4 What can be done in near future?

The information can be put in specific format as described earlier which will certainly increase efficiency. Many websites nowadays are developing directories which already have the categorized webpages. If some additional effort is taken then we can even categorize these existing webpages according to their accessibility to various users. Some of the webpages have already such facility of choosing your level Mathematics expertise before searching for the information, [4]. Of course it is not possible to put all the available information in the above mentioned format as it is a naïve categorization.

4 Need for categorization

We have already discussed the problem at length to show the need and how our approach can help to solve the problem. There is one more additional reason for this. The users when searching for some information, hardly check first two or three webpages, [1]. This even forces us more to present the information in the proper format. Also in the proposed format the user can scan through the information in less time than in usual jumbled up format. Representing our idea in concise format, that is what one strives while writing abstract for the paper [2], then why not apply the idea to all the knowledge base.

5 Future work

We would like to develop a small knowledge base using the our approach. Then we will possibly put it on the web to get the feedback from users worldwide. As already mentioned there are some websites which are trying to do this so we will try to contact them to get their views on the problem and to know the response to their solution.

Our approach is naïve and have to be modified for large set of information. Some more categories like teacher need to be introduced and narrower classification might help.

References

[1] P. Bailey, N. Craswell, K. Griffiths, D. Hawking. **Measuring Search Engine Quality.** *Information Retrieval*, Vol. 4, No. 1, pp. 33–59, April 2001

[2] B. Buchberger. **Thinking, Speaking and Writing.** *Lecture notes*, 1995.

[3] `http://forum.swarthmore.edu/dr.math/`

[4] `http://nrich.maths.org/mathsf/links/index.php3`

[5] `http://www1.AskMe.com/BusinessSplash.asp`

[6] `http://www.bartleby.com/110/3000.html`

[7] `http://www.gc.cc.fl.us/leolusk/divisibi.htm`

[8] `http://www.math-atlas.org/`

[9] `http://www2.sunysuffolk.edu/wrightj/MA27/PolyExp/Divide.htm`

[10] E. Kaltofen. **Challenges of Symbolic Computation: My Favorite Open Problems.** *Journal of Symbolic Computation*, Vol. 29, No. 6, pp. 891–919, June 2000.

[11] S. Lang. **The Beauty of Doing Mathematics : Three Public Dialogues.** Springer, New York, 1985.