Solution of Algebraic Equations by Using Autonomous Computational Methods

Andrew Pownuk[1], Jose Gonzalez[2]

[1]Department of Mathematical Sciences, University of Texas at El Paso, El Paso, Texas, ampownuk@utep.edu
[2]Undergraduate Student at the Department of Electrical and Computer Engineering, University of Texas at El Paso, El Paso, Texas, jhgonzalez5@miners.utep.edu

25th Joint NMSU/UTEP Workshop on Mathematics, Computer Science, and Computational Sciences

Outline

Checkers (Game Theory)

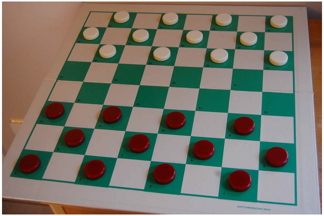


Figure: Game tree

Arthur Samuel in 1959 applied alpha–beta pruning in order to create a program that learned how to play checkers. Alpha–beta pruning is a search algorithm that seeks to decrease the number of nodes that are evaluated by the minimax algorithm in its search tree.

Game Theory (tic-tac-toe)

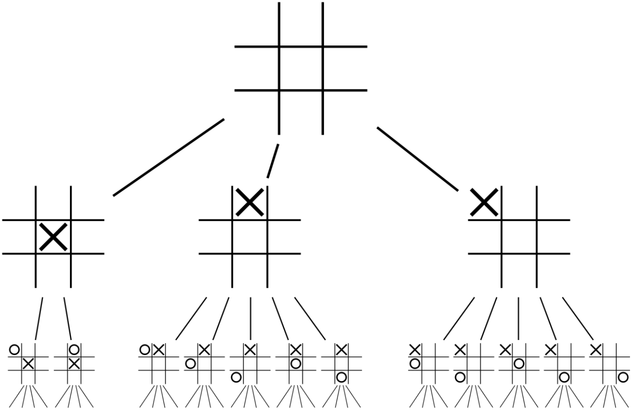


Figure: Game tree

Game Theory (GO)

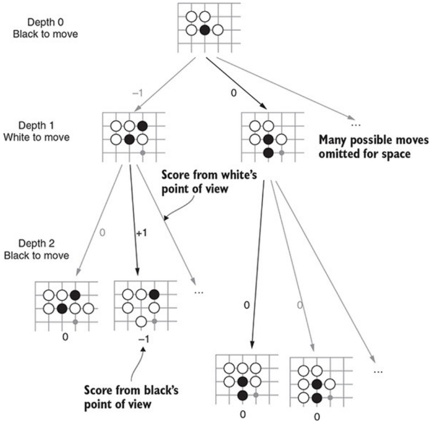


Figure: Game tree

Game Theory (winning strategy)

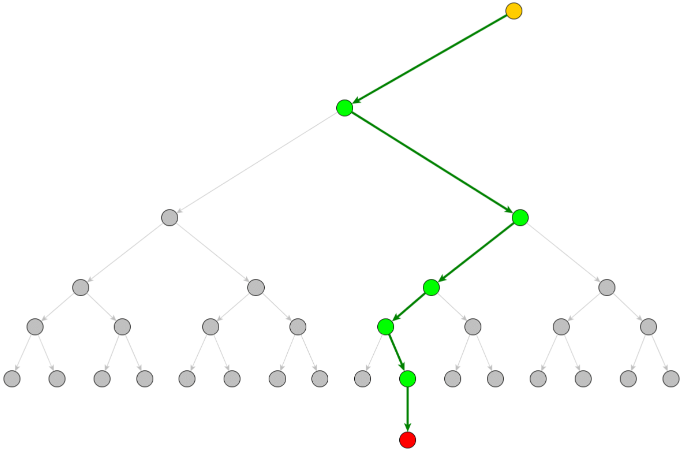


Figure: Game tree

Basic Mathematical Operations

Addition

$$1 + 1 = 2$$

$$1 + 0 = 1$$

$$2 + (-1) = 1$$

Multiplication

$$1 * 0 = 0$$

$$i * i = -1$$

Algebra

$$a + b = b + a$$

$$a(b + c) = ab + ac$$

etc.

Differentiation

Input information

- $(f(x)+g(x))' = f'(x)+g'(x)$
- $(f(x)-g(x))' = f'(x)-g'(x)$
- $(f(x)g(x))' = f'(x)g(x)+f(x)g'(x)$
- $(f(x)/g(x))' = \frac{f'(x)g(x)-f(x)g'(x)}{g^2(x)}$
- $(f(g(x)))' = f'(g(x))g'(x)$
- arithmetic operations

Differentiation - Sample Application

Product rule (input information)

$$(f * g)' = f' * g + f * g'$$

After calculations (new theorem created automatically)

$$((f * g) * h)' = (f * g)' * h + (f * g) * h'$$

$$((f * g) * h)' = (f' * g + f * g') * h + (f * g) * h'$$
$$(f * g * h)' = (f' * g) * h + (f * g') * h + (f * g) * h'$$

New theorem can be used in exactly the same way like the original theorem.

$$(f * g * h)' = f' * g * h + f * g' * h + f * g * h'$$

Differentiation (Latex source)


```
\begin{equation}\cos\left(x\right)+x+2+\sin\left(2\right)\end{equation}
\begin{equation}\cos\left(x\right)+x+2+\sin\left(x\right)\end{equation}
\begin{equation}\cos\left(x\right)+x+2+\sin\left(\sin\left(x\right)\right)\end{equation}
\begin{equation}\cos\left(x\right)+x+2+\sin\left(\cos\left(x\right)\right)\end{equation}
\begin{equation}\cos\left(x\right)+x+2+\cos\left(2\right)\end{equation}
\begin{equation}\cos\left(x\right)+x+2+\cos\left(x\right)\end{equation}
\begin{equation}\cos\left(x\right)+x+2+\cos\left(\sin\left(x\right)\right)\end{equation}
\begin{equation}\cos\left(x\right)+x+2+\cos\left(\cos\left(x\right)\right)\end{equation}
\begin{equation}\cos\left(x\right)+x+2+2+2\end{equation}
\begin{equation}\cos\left(x\right)+x+2+2+x\end{equation}
\begin{equation}\cos\left(x\right)+x+2+2+\sin\left(x\right)\end{equation}
\begin{equation}\cos\left(x\right)+x+2+2+\cos\left(x\right)\end{equation}
\begin{equation}\cos\left(x\right)+x+2+x+x\end{equation}
\begin{equation}\cos\left(x\right)+x+2+x+\sin\left(x\right)\end{equation}
\begin{equation}\cos\left(x\right)+x+2+x+\cos\left(x\right)\end{equation}
\begin{equation}\cos\left(x\right)+x+2+\sin\left(x\right)+\sin\left(x\right)\end{equation}
\begin{equation}\cos\left(x\right)+x+2+\sin\left(x\right)+\cos\left(x\right)\end{equation}
\begin{equation}\cos\left(x\right)+x+2+\cos\left(x\right)+\cos\left(x\right)\end{equation}
\begin{equation}\cos\left(x\right)+x+x+x\end{equation}
```

Differentiation (step 1)

$$\cos(x) + (\cos(x) + \cos(2)) \tag{765}$$
$$\cos(x) + (\cos(x) + \cos(x)) \tag{766}$$
$$\cos(x) + (\cos(x) + \cos(\sin(x))) \tag{767}$$
$$\cos(x) + (\cos(x) + \cos(\cos(x))) \tag{768}$$
$$\cos(x) + (\cos(x) + (2 + 2)) \tag{769}$$
$$\cos(x) + (\cos(x) + (2 + x)) \tag{770}$$
$$\cos(x) + (\cos(x) + (2 + \sin(x))) \tag{771}$$
$$\cos(x) + (\cos(x) + (2 + \cos(x))) \tag{772}$$
$$\cos(x) + (\cos(x) + (x + x)) \tag{773}$$

Differentiation (step 2)

$$\frac{d}{dx}\cos(x) + \frac{d}{dx}(\cos(x) + \cos(2)) \tag{765}$$

$$\frac{d}{dx}\cos(x) + \frac{d}{dx}(\cos(x) + \cos(x)) \tag{766}$$

$$\frac{d}{dx}\cos(x) + \frac{d}{dx}(\cos(x) + \cos(\sin(x))) \tag{767}$$

$$\frac{d}{dx}\cos(x) + \frac{d}{dx}(\cos(x) + \cos(\cos(x))) \tag{768}$$

$$\frac{d}{dx}\cos(x) + \frac{d}{dx}(\cos(x) + (2 + 2)) \tag{769}$$

$$\frac{d}{dx}\cos(x) + \frac{d}{dx}(\cos(x) + (2 + x)) \tag{770}$$

$$\frac{d}{dx}\cos(x) + \frac{d}{dx}(\cos(x) + (2 + \sin(x))) \tag{771}$$

Differentiation (step 3)

$$(-1) \cdot \sin(x) + \left(\frac{d}{dx} \cos(x) + \frac{d}{dx} \sin(x) \right) \tag{762}$$

$$(-1) \cdot \sin(x) + \left(\frac{d}{dx} \cos(x) + \frac{d}{dx} \sin(\sin(x)) \right) \tag{763}$$

$$(-1) \cdot \sin(x) + \left(\frac{d}{dx} \cos(x) + \frac{d}{dx} \sin(\cos(x)) \right) \tag{764}$$

$$(-1) \cdot \sin(x) + \left(\frac{d}{dx} \cos(x) + \frac{d}{dx} \cos(2) \right) \tag{765}$$

$$(-1) \cdot \sin(x) + \left(\frac{d}{dx} \cos(x) + \frac{d}{dx} \cos(x) \right) \tag{766}$$

$$(-1) \cdot \sin(x) + \left(\frac{d}{dx} \cos(x) + \frac{d}{dx} \cos(\sin(x)) \right) \tag{767}$$

Differentiation (step 4)

$$(-1)\cdot\sin(x)+\left((-1)\cdot\sin(x)+\left(\frac{d}{du}\cos(u)\right)\Bigg|_{u=\sin(x)}\cdot\frac{d}{dx}\sin(x)\right) \tag{767}$$

$$(-1)\cdot\sin(x)+\left((-1)\cdot\sin(x)+\left(\frac{d}{du}\cos(u)\right)\Bigg|_{u=\cos(x)}\cdot\frac{d}{dx}\cos(x)\right) \tag{768}$$

$$(-1)\cdot\sin(x)+\left((-1)\cdot\sin(x)+\left(\frac{d}{dx}2+\frac{d}{dx}2\right)\right) \tag{769}$$

$$(-1)\cdot\sin(x)+\left((-1)\cdot\sin(x)+\left(\frac{d}{dx}2+\frac{d}{dx}x\right)\right) \tag{770}$$

$$(-1)\cdot\sin(x)+\left((-1)\cdot\sin(x)+\left(\frac{d}{dx}2+\frac{d}{dx}\sin(x)\right)\right) \tag{771}$$

$$(-1)\cdot\sin(x)+\left((-1)\cdot\sin(x)+\left(\frac{d}{dx}2+\frac{d}{dx}\cos(x)\right)\right) \tag{772}$$

$$(-1)\cdot\sin(x)+\left((-1)\cdot\sin(x)+\left(\frac{d}{dx}x+\frac{d}{dx}x\right)\right) \tag{773}$$

Differentiation (step 5)

$$(-1)\cdot\sin(x)+((-1)\cdot\sin(x)+\cos(\sin(x))\cdot\cos(x)) \tag{763}$$
$$(-1)\cdot\sin(x)+((-1)\cdot\sin(x)+\cos(\cos(x))\cdot(-1)\cdot\sin(x)) \tag{764}$$
$$(-1)\cdot\sin(x)+((-1)\cdot\sin(x)+0\cdot 0) \tag{765}$$
$$(-1)\cdot\sin(x)+((-1)\cdot\sin(x)+(-1)\cdot\sin(x)) \tag{766}$$
$$(-1)\cdot\sin(x)+((-1)\cdot\sin(x)+(-1)\cdot\sin(\sin(x))\cdot\cos(x)) \tag{767}$$
$$(-1)\cdot\sin(x)+((-1)\cdot\sin(x)+(-1)\cdot\sin(\cos(x))\cdot(-1)\cdot\sin(x)) \tag{768}$$
$$(-1)\cdot\sin(x)+((-1)\cdot\sin(x)+(0+0)) \tag{769}$$
$$(-1)\cdot\sin(x)+((-1)\cdot\sin(x)+(0+1)) \tag{770}$$
$$(-1)\cdot\sin(x)+((-1)\cdot\sin(x)+(0+\cos(x))) \tag{771}$$
$$(-1)\cdot\sin(x)+((-1)\cdot\sin(x)+(0+(-1)\cdot\sin(x))) \tag{772}$$
$$(-1)\cdot\sin(x)+((-1)\cdot\sin(x)+(1+1)) \tag{773}$$

Group Axioms

A group is a set (mathematics), G together with an Binary operation $''\cdot''$ (called the group law of G) that combines any two element elements a and b to form another element, denoted $a \cdot b$. To qualify as a group, the set and operation, $(G, \cdot)$, must satisfy four requirements known as the group axioms:

1. Closure: For all $a, b \in G$, the result of the operation, $a \cdot b$, is also in G.
2. Associativity: For all a, b and c in G, $(a \cdot b) \cdot c = a \cdot (b \cdot c)$.
3. Identity element: There exists an element e in G such that, for every element a in G, the equation $1 = e \cdot a = a \cdot e = a$ holds. Such an element is unique.
4. Inverse element: For each $a \in G$, there exists an element $b \in G$, commonly denoted a^{-1} (or $-a$, if the operation is denoted $+$, such that $a \cdot b = b \cdot a = e$, where e is the identity element.

Field Axioms

1. Addition is commutative: $x + y = y + x$, for all $x, y \in F$.
2. Addition is associative: $(x + y) + z = x + (y + z)$, for all $x, y, z \in F$.
3. Existence of additive identity: there is a unique element $0 \in F$ such that $x + 0 = x$, for all $x \in F$.
4. Existence of additive inverses: if $x \in F$, there is a unique element $-x \in F$ such that $x + (-x) = 0$.
5. Multiplication is commutative: $xy = yx$, for all $x, y \in F$.
6. Multiplication is associative: $(xy)z = x(yz)$, for all $x, y, z \in F$.
7. Existence of multliplicative identity: there is a unique element $1 \in F$ such that $1 \neq 0$ and $x1 = x$, for all $x \in F$.
8. Existence of multliplicative inverses: if $x \in F$ and $xneq0$, there is a unique element $(1/x) \in F$ such that $x(1/x) = 1$.
9. Distributivity: $x(y + z) = xy + xz$, for all $x, y, z \in F$.

Expression Evaluation

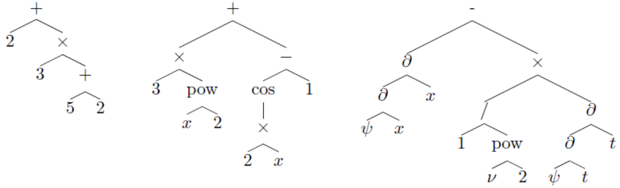
step-1
$(((1+2)+((1+2)-(3+4)))+(5+(6+7)))=((18-1)+0)$
step-2
$(((1+2)+((1+2)-(3+4)))+(5+(6+7)))=(17+0)$
step-3
$((3+((1+2)-(3+4)))+(5+(6+7)))=((18-1)+0)$
step-4
$(((1+2)+((1+2)-(3+4)))+(5+13))=((18-1)+0)$
step-5
$(((1+2)+(3-(3+4)))+(5+(6+7)))=((18-1)+0)$
step-6
$(((1+2)+((1+2)-7))+(5+(6+7)))=((18-1)+0)$
step-7
$(((1+2)+((1+2)-(3+4)))+(5+(6+7)))=17$
step-8
$((3+((1+2)-(3+4)))+(5+(6+7)))=(17+0)$
step-9
$(((1+2)+((1+2)-(3+4)))+(5+13))=(17+0)$
step-10
$(((1+2)+(3-(3+4)))+(5+(6+7)))=(17+0)$

Expression Evaluation

step-93
$((3 + (-4)) + 18) = (17 + 0)$
step-94
$((-1) + (5 + 13)) = (17 + 0)$
step-95
$((-1) + 18) = ((18 - 1) + 0)$
step-96
$((3 + (-4)) + 18) = 17$
step-97
$((-1) + (5 + 13)) = 17$
step-98
$((-1) + 18) = (17 + 0)$
step-99
$17 = ((18 - 1) + 0)$
step-100
$((-1) + 18) = 17$
step-101
$17 = (17 + 0)$
step-102
$17 = 17$

Expression evaluation is possible without specifying any explicit method for expression evaluation.

Computational Graph



COCONUT Project
COntinuous CONstraints - Updating the Technology.
https://www.mat.univie.ac.at/ neum/glopt/coconut/

Solution of Algebraic Equations

step-1

$x = (x + (2 + x))$

step-3

$x = (x + 2 + x)$

step-4

$x = (x + 2 + x)$

step-5

$x = x + 2 + x$

step-6

$x = x + 2 + x$

step-7

$x = 2 + 2 * x$

step-9

$x + (-1) * x + (-1) * x = 2$

step-10

$x + (-2) * x = 2$

step-11

$(-1) * x = 2$

step-12

$x = (2/(-1))$

step-13

$x = (-2)$

step-14

$x = ((-2)/1)$

Automatically Generated Latex Report


```
step-1\\
$x=(x+(2+x))$
\newline
step-3\\
$x=(x+2+x)$
\newline
step-4\\
$x=(x+2+x)$
\newline
step-5\\
$x=x+2+x$
\newline
step-6\\
$x=x+2+x$
\newline
step-7\\
$x=2+2*x$
\newline
step-9\\
$x+(-1)*x+(-1)*x=2$
```

Automatically Generated PDF Reports

Hundred/thousand/unlimited number of examples:

51810	data-node-126.pdf
52690	data-node-127.pdf
52354	data-node-128.pdf
52785	data-node-130.pdf
52603	data-node-131.pdf
52512	data-node-132.pdf
52392	data-node-134.pdf
52565	data-node-135.pdf
53310	data-node-136.pdf
50225	data-node-14.pdf
52728	data-node-142.pdf
51848	data-node-143.pdf
52303	data-node-144.pdf
52666	data-node-146.pdf
52696	data-node-147.pdf
52484	data-node-148.pdf
50286	data-node-15.pdf
52539	data-node-150.pdf

http://andrew.pownuk.com/research/AlgebraicEquations/

Simplification of the Solution

step-1
$(4+6) = (x+x)$
step-3
$4+6 = (x+x)$
step-7
$4+6 = x+x$
step-11
$4+6 = 2*x$
step-15
$4+6+(-2)*x = 0$
step-21
$(-2)*x = 0+(-1)*4+(-1)*6$
step-24
$(-2)*x = (-10)$
step-25
$x = ((-10)/(-2))$
step-26
$x = 5$

Simplification of the Solution

- Finding optimal form of the computational process is one or the main goal of this project.
- Quality of the simplification depend on the amount of knowledge available in the system and processing power.
- For small problem the simplification can be found uniquely.
- For more complex problem simplification of the expressions never stops.
- It is possible to use new computational results in order future calculations (self-adaptivity). In this way the system can generate knowledge in autonomous way.

Finite Number of Steps

- Limits $\lim_{n\to\infty} \frac{1}{n} = 0$
- Convergence of infinite series $\sum_{n=1}^{\infty} \frac{1}{n^2}$ ($p = 2 > 1$ series converges).
- If function f is integrable, then the sequence of Riemann sum $\sum_{n=1}^{N} f(x_n)\Delta x_n$ converges to $\int_a^b f(x)dx$ (for appropriate partitions).
- If function f is integrable, then the sequence of Riemann sum $\sum_{n=1}^{N} f(x_n)\Delta x_n$ converges to $\int_a^b f(x)dx$ (for appropriate partitions).

Presented methodology can be applied to many scientific theories (mathematics, engineering, chemistry, biology etc.) with finite number of steps.

What is Special About This Research?

All calculations are done in fully **autonomous** way.

Why autonomous calculations?

What is Special About This Research?

All calculations are done in fully **autonomous** way.

Why autonomous calculations?

Correctness
and
Scalability.

What is Special About This Research?

All calculations are done in fully **autonomous** way.

Why autonomous calculations?

Results are NOT Biased and Subjective .

Classification problem

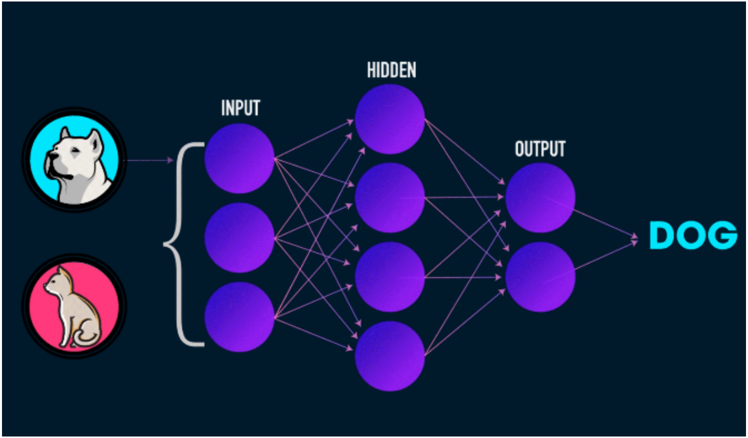
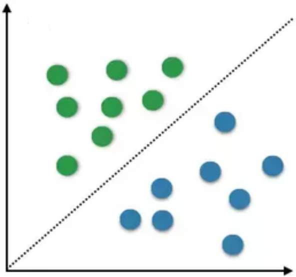


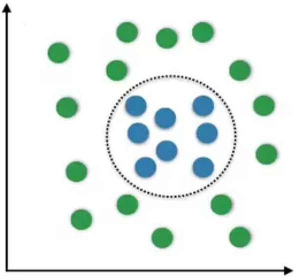
Figure: Classification of pictures

Classification problem from mathematical point of view

Example

If x_1 is in the class A then $f(x_1, W) > 0$.
If x_2 is in the class B then $f(x_2, W) < 0$.





Problem space

- Trees with up to n internal nodes.
- Set p_1 of unary operators (e.g. cos, sin, exp, log).
- Set p_2 of binary operators (e.g. $+$, $-$, $/$, $\times$).
- a set of L leaf values containing variables (e.g. x, y, z), constants (e.g. e, π), integers (e.g. $\{$,-10,..., 10,$\}$).

The number of binary trees with n internal nodes is given by the n-th Catalan numbers.
A binary tree with n internal nodes has exactly $n+1$ leaves. Each node and leaf can take respectively p_2 and L different values. As a result, the number of expressions with n binary operators can be expressed by:

$$E_n = \frac{1}{n+1}\binom{2n}{n} p_n^n L^{n+1} \approx \frac{4^n}{n\sqrt{\pi \cdot n}}(p_2)^n L^{n+1}$$

Training of the Model

Least square error:

$$J(W) = \sum_{i=1}^{n} \sum_{j=1}^{n_O} \left(y_j^{(i)} - \hat{y}_j^{(i)} \right)^2$$

Learning process

$$W^* = \arg\min J(W) = \arg\min \sum_{i=1}^{n} \sum_{j=1}^{n_O} \left(y_j^{(i)} - \hat{y}_j^{(i)} \right)^2$$

Prediction

$$y_j = \hat{y}_j(x, W^*)$$

Predicting Solution Method

Let x is a given mathematical problem, then after the training it is possible to predict the solution step y.

$$y = f(x, W)$$

Google AI system proves over 1200 mathematical theorems, April 26th, 2019.

Kshitij Bansal, Sarah M Loos, Markus N. Rabe, Christian Szegedy, and Stewart Wilcox, HOList: An Environment for Machine Learning of Higher-Order Theorem Proving, 24 May 2019, https://arxiv.org/pdf/1904.03241.pdf

Finite Number of Steps? Logic?

- The fundamental difference between if-else and AI is that if-else models are static deterministic environments and machine learning (ML) algorithms, the primary underpinning of AI, are probabilistic stochastic environments.
- Remark: machine learning (probabilistic) reasoning use mathematics which is based on above described methodology.
- Cyc - is the world's longest-lived artificial intelligence project, attempting to assemble a comprehensive ontology and knowledge base. Douglas Lenat began the project in July 1984 at MCC, where he was Principal Scientist 1984–1994, and then, since January 1995, has been under active development by the Cycorp company, where he is the CEO.

Solution of Equations with Uncertain Parameters

The main objective of my present research is to build an autonomous system for automated development of scientific knowledge. The system will be/is applied to automated development of scientific theory of equations with uncertain parameters. The system will be/is capable to automatically expand new scientific ideas (on the basis of existing background knowledge) as well as to improve itself. System would gather existing knowledge, check it in new selected directions, document the process and results, and save new algorithms generated in the process. Once performed research and generated results will be remembered in the system and possible to use if needed.

Computer Based Research Tools

Scientists use many methods and techniques to perform their research. Also supporting tools are necessary elements for research performance. Discovery/research supporting tools can be as simple as piece of paper and a pencil, but in most subjects they are much more complicated, and recently except specific apparatus support is mostly delivered by computers and advanced software. The computer based research tools should support mathematicians, scientists, and engineers help them make connections to related fields.

Conclusions

- Mathematical/scientific knowledge can be treated as independent units that can interact with each-other and create new, possibly useful knowledge.
- Generation of new knowledge can be fully automated and autonomous. No interaction with humans is necessary.
- Development of new knowledge is possible in many different fields (e.g. statistics, engineering, chemistry, biology, computer science etc.).

Conclusions (continued)

- By using presented methodology it is possible to create complex examples and appropriate computational methods relevant to many areas of mathematics as well as in other areas of science and engineering.
- Scientific results can be created in fully objective way without biased opinions of human researchers.
- By using self-adaptive computational methods it is possible to automatically generate new mathematical theorems without interactions with humans (consequently without human errors).
- Machine learning (the main computational method is NOT based on machine learning) can be used as source of good initial guess for processing mathematical information. Actually there is NO main computational method in the system.

Conclusions (continued)

- Once the information is available in the system it will NEVER be forgotten and can be used for generation of new mathematical theorems. From that perspective the system can be viewed as self-organizing archive of information.
- Development of this and similar systems should speed up cooperation among scientists around the world (theoretical possibility).

Conclusions (continued)

- Calculations can be done in distributed way (this option is experimental at this moment). Unlimited number of computers can process simultaneously in order to get the results faster. Calculations do not require existence of any centralized system.
- Turning off some computers slows down the calculation.
- Parallel computing can significantly speed up the calculations (future work).
- Autonomous interaction with external sources of information extends internal database of information and should increase productivity of the system (future work).

Conclusions (continued)

- Mathematical theorems can be used for the solution of many practical engineering and scientific problems if appropriate domain specific knowledge is available.
- I created some practical examples (civil engineering, oil engineering problems) by using presented system but not in fully autonomous way (work in progress).

Thank you