3. Многонишково програмиране (Scalable Multithreaded Programming) с "Thread Pools" (паралелна сортировка)

Съществуват много начини за разпределяне задачите по процесорни ядра или в многопроцесорна среда.

Ще преобразуваме еднонишково приложение към такова, ползващо всички налични изчислителни ресурси.

Ще въведем основни понятия от:

- OpenMP технологията и
- --thread pools.

Ползвайки Visual Studio utilities, ще демонстрираме измерване на подобренията в производителността.

В началото, ще раздробим общата задача на по-малки, подходящи за нишково разпределяне подзадачи, които бихме могли да подадем към отделните ядра.

Спазваме няколко препоръки при това преобразуване:

- Задачата да подлежи на разпаралеляване.
- Подзадачите да са независими помежду си.
- Да могат да се изпълняват в произволен ред.
- Да поддържат собствени копия на данните си.

3.1 Multithreading c OpenMP

OpenMP е една от най-простите технологии за въвеждане на паралелизъм и се поддържа от Visual Studio C++ компилатора след версия 2005. Въвеждането на паралелизма става чрез "OpenMP - <u>pragmas"</u> директивите в кода. Пример:

директивата 'pragma' паралелизира следващия я блок код в случая това е само printf()— и го пуска едновременно по всички изчислителни ресурса.

Броят им варира - според инсталираните процесори или ядра.

За да разрешите паралелизацията с OpenMP (компилаторът да не игнорира "pragma" директивите), следва да разрешите OpenMP. За целта: първо

опция към компилатора /openmp

(Properties / C/C++ / Language / Open MP Support). Второ: include the omp.h

Следва пример за разпаралеляване на матрични изчисления (повдигане на степен) и обработка върху всички налични ядра:

```
#pragma omp parallel for
for (int i = 0; i < 50000; i++)
array[i] = i * i;</pre>
```

OpenMP притежава и конструктори за контрол на:

- броя създавани нишки;
- Управление последователността на отделните запаралелени блокове;
- Създаване на thread-local data,
 - точки на синхронизация,
 - критични секции и др.

OpenMP е лесен начин за въвеждане на паралелизъм в съществуващ вече код.

OpenMP е проста технология. В много случаи е необходим по-голям контрол над изчислителния процес.

Например, разпределяне нишките по конкретни ядра, динамичен паралелизъм и т.н..

Идват други технологични решения –например, thread-pools.

3.2 Thread Pool

Нишките се менажират от OS. Те заемат ресурс и създават свое обкръжение. Следователно, честото им създаване и унищожаване е скъпа операция. По-добре е вече създадените да отлежават за повторно използване, при необходимост.

Това място за изчакване се нарича thread pool и Windows поема поддръжката му

Ползването на thread pool сваля от програмиста отговорността за често създаване на нишки, унищожаването и управлението им.



Удобно и полезно е да се разпаралели задача в много ядра.

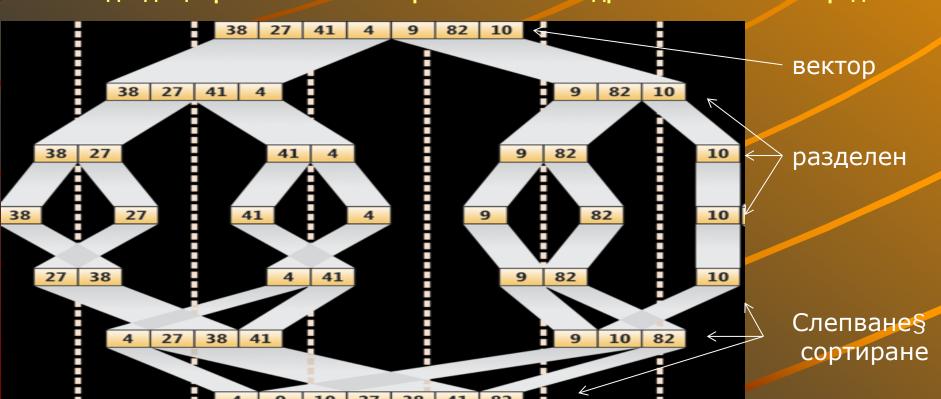
3.3 пример с многонишково сортиране

Сортировката не е най-силния алгоритъм за паралелизация. Тук има пречка – как да разпаралеляваме, така че да подсигурим независимост на отделните блокове?

Наивен подход е да заключим достъпа до данните за времето на обработка например чрез mutex, semaphore или критична секция.

По-добро решение е на всяко ядро да се подаде подсекция на масива данни за паралелна сортировка. Този divide-and-conquer подход е като че ли по-подходящ.

Алгоритмите merge sort и quick sort работят добре с тази стратегия и тя е подходящо решение за натоварване на многоядрена изчислителна среда.



Разделените подсписъци са независими и могат да се подадат към СРU ядрата за паралелна обработка без заключвания.

Съществуват много алгоритми за сортировка, податливи на Паралелизация:

- Quick sort,
- selection sort,
- merge sort,
- radix sort

Всички те разбиват данните и ги обработват независимо.

Примерна реализация на сортировка е показана (quick sort на C#).

Програмата инициализира голям масив с поредица случайни числа и после ги сортира с помощта на quick sort процедура, като отчита и времето за това.

```
2
```

```
namespace ParallelSort {
 class Program {
  // For small arrays, use Insertion Sort
  private static void InsertionSort(
             int[] list, int left, int right)
  for (int i = left; i < right; i++)
       int temp = list[i];
      int j = i;
       while ((j > 0) \&\& (list[j - 1] > temp))
           { list[j] = list[j - 1];
              i = i - 1;
       list[j] = temp;
private static int Partition( int[] array, int i, int j)
     int pivot = array[i];
    while (i < j) {
         while (array[j] >= pivot && i < j)
        while (array[i] <= pivot && i < j)
                       \{i++:
        if (i < j) { array[j--] = array[i]; }
            array[i] = pivot;
     return i;
```

```
static void QuickSort( int[] array, int left, int right)
 { // Single or 0 elements are already sorted
   if (left >= right)
                         return:
  // For small arrays, use a faster serial routine
   if ( right-left <= 32) { InsertionSort(array, left, right);</pre>
                      return:
  // Select a pivot, then quicksort each sub-array
   int pivot = Partition(array, left, right);
   QuickSort(array, left, pivot - 1);
   QuickSort(array, pivot + 1, right);
  static void Main(string[] args)
   const int ArraySize = 50000000;
   for (int iters = 0; iters < 1; iters++)
      { int[] array;
        Stopwatch stopwatch:
        array = new int[ArraySize];
        Random\ random1 = new\ Random(5);
       for (int i = 0; i < array.Length; ++i)
          { array[i] = random1.Next(); }
       stopwatch = Stopwatch.StartNew();
       QuickSort(array, 0, array.Length - 1);
       stopwatch.Stop();
       Console.WriteLine("Serialt: {0} ms",
      stopwatch.ElapsedMilliseconds);
```

}}

За да паралелизирате приложението, променете следните оператори:

```
QuickSort( array, lo, pivot - 1);
QuickSort( array, pivot + 1, hi);
```

Към паралелната реализация:

```
Parallel.Invoke(
```

```
delegate { QuickSort(array, left, pivot - 1); },
delegate { QuickSort(array, pivot + 1, right); }
);
```

Интерфейсът Parallel Invoke от Systems. Threading. Tasks namespace е дефиниан в .NET Task Parallel Library. Той позволява дефиниране на асинхронно изпълнявана функция. В нашата реализация - всяка отделна сортировъчна функция ще се изпълни в отделна нишка.

Макар че е по-добре да запуснете само 1 нишка, а втория подсписък да се поеме за сортиране от текущата, реализация с 2 запуснати нишки за сортировка е по-симетрична и илюстрира колко лесно е преобразуването на серийна програма в паралелен еквивалент.

Разумен въпрос е: паралелизацията подобри ли производителността?

Visual Studio 2010 включва няколко tools за такава оценка :

Visual Studio 2010 Performance Wizard мери производителност в паралелна реализация по време на изпълнение.

Можете да наблюдавате и core utilization.

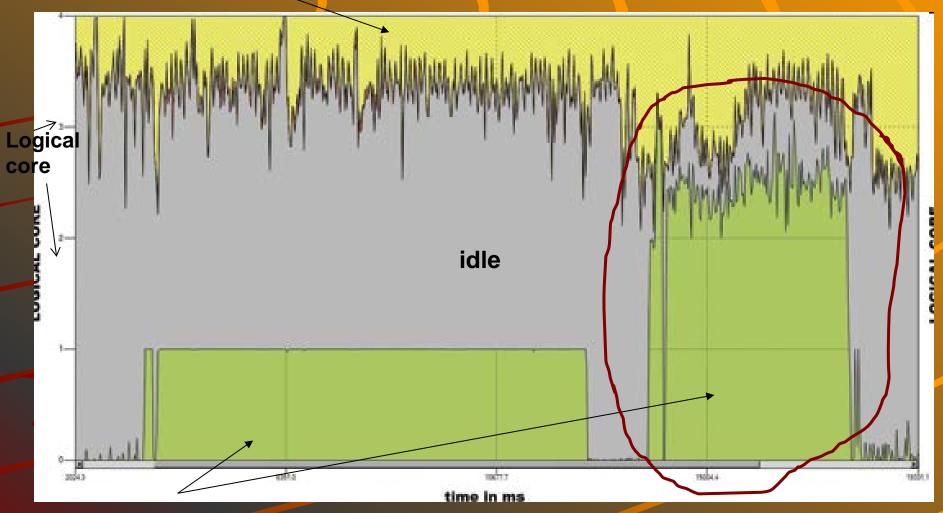
Нека тестовата програма стартира серийна сортировка, заспива секунда и тогава стартира паралелна версия на сортировъчния алгоритъм.

За 4-ядрен компютър, получваме следната графична представа:

В началото е видно натоварването при 1 ядрена реализация – 100% използваемост на единственото ядро. Впоследствие се вижда 2.25 кратно ускорение.

Паралелната реализация ускорява с около 45% в сравнение със серийния си еквивалент.

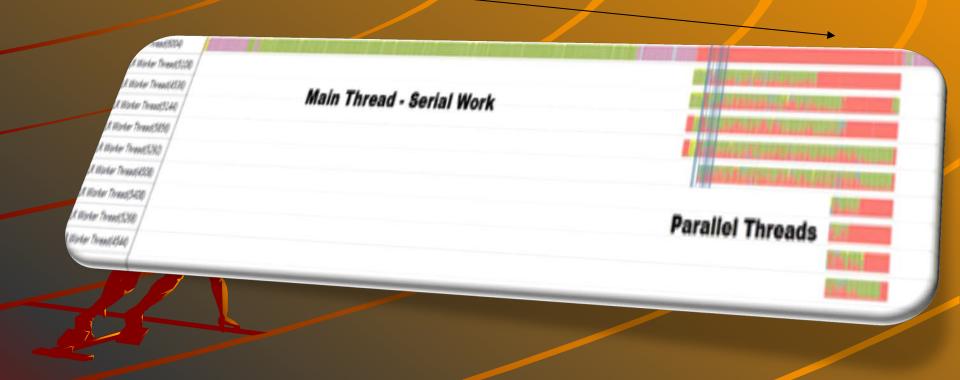
OS и други програми



приложението

Друга визуализация

Показваме как приложението ползва наличните нишки. Една нишка работи почти цялото време. Розовият цвят показва нишки, които са блокирани от други.



Вижда се, че макар натоварването на СРU ядра е подобрено, има потенциал за още значителни оптимизации.

4. Динамично програмиране

Рекурсията не е оптималната техника →в ред случаи рекурсивният алгоритъм следва а се пренапише в нерекурсивен вариант със запомняне междинни резултати в таблица.

Това е "динамичното програмиране".

4.1 таблици вместо рекурсия

Ето неефективен вариант за изчисляване числата на Фибуначи:

```
int fib( in n)

    if( n <= 1 )
        return 1;
    else
    return fib( n - 1 ) + fib( n - 2 );
}</pre>
```

$$T(N) >= T(N-1) + T(N-2)$$

Вече изследвана в курса сложност: експоненциална

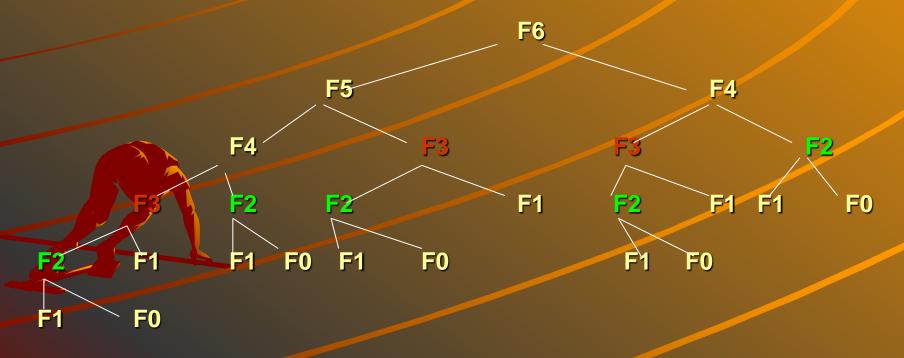
Можем, обаче, да съхраним изчислените вече F n-1; F n-2.:

```
int fibonacci (int n )
        if( n <= 1 )
               return 1;
        int last = 1;
        int nextTolast = 1;
        int \ answer = 1;
        for(int i = 2; i <= n; i++)
                answer = last + nextTolast;
                nextTolast = last;
                last = answer;
        return answer;
```

Имаме O(N).

Ако горната модификация не е направена, за изчисление на Fn вика рекурсивно Fn-1 и Fn-2.....

Ето трасировка на рекурсиите.



Виждаме F N-3 се вика 3 пъти, FN-4 – 5 пъти и т.н.

Експлозивно нарастват сметките

4.2 оптимално бинарно търсене в дърво

Имаме списък думи w1......wn с вероятности на появяване: p1.....pn Искаме да подредим думите в бинарно дърво, така че общото време за намиране на дума да минимално.

В бинарно дърво, за да стигнем до елемент с дълбочина d, правим d+1 сравнения. Така че, ако w_i е на дълбочина d_i, ние искаме да минимизираме <u>n</u>

pi (1+di). дума вероятност 0.22 0.18 am two 0.20 and the 0.05 egg and if 0.25 Някаква Greedy стратегия: the 0.02 **^ вероатност** → **ляв** клон ám egg 80.0 two (and) egg am two and two a am egg Друга подредба: балансирано, бинарно д-во Оптимално, the но неинтуитивно д-во

първото използва greedy стратегия. Най-вероятната дума е най-горе. Второто е балансирано search tree (в дясно с по-ниско рі). Третото е оптималното – виж таблицата:

вход		дърво	1	дърво2	2		дърво3
дума	вероятнос	т	цена		цена		цена
		дълбочина	ДЪ	лбочина	дъ	лбочина	
a	0.22	2	0.44	3	0.66	2	0.44
am	0.18	4	0.72	2	0.36	3	0.54
and	0.20	3	0.60	3	0.60	1	0.20
egg	0.05	4	0.20	1	0.05	3	0.15
if	0.25	1	0.25	3	0.75	2	0.50
the	0.02	3	0.06	2	0.04	4	0.08
two	0.08	2	0.16	3	0.24	3	0.24
общо	1.00		2.43		2.70		2.15

сравнение на 3 дървета с бинарно търсене

При оптимално, бинарно "search tree" (формирано по алгоритъм, подобен на този на Huffman) данните не са само в листата (както беше при Huffman компрерсията) и следва да удовлетворяваме критерий за binary search tree.

Поставяме сортирани според някакъв критерий думи Wleft, Wleft+1,....Wright-1, Wright в бинарно дърво. Нека сме постигнали оптималното бинарно дърво в което имаме корен Wi и под-дървета за които: Left <= i <= Right

Тогава лявото поддърво съдържа елементите Wieft, ...W i-1, а дясното – Wi+1,.....Wright (критерий за binary search tree).

Поддърветата също правим оптимални. Тогава можем да напишем формулата за цената Cleft, right на оптимално binary search tree. Лявото поддърво има цена Cleft, I-1, дясното – Ci+1, right, спрямо корена си.



Структура на оптимално за претърсване, бинарно дърво

търсим минимизация за цената на цялото дърво (т.е. кое "і" е в корена):

Cleft,Right = min
$$pi + CLeft,i-1 + C i+1,Right + \sum_{j=Left} pj + \sum_{j=i+1} Right$$
 Спрямо това і $pi + CLeft,i-1 + Ci+1,Right + \sum_{j=Left} pj + \sum_{j=i+1} Right$ За да отчетем че всеки възел в тези поддървета е 1 ниво по-дълбоко спрямо общия корен

На основата на тази формула се гради алгоритъмът за минимална цена (оптималното дърво).

Търсим і, така че да се минимизира С Left, Right.

Таблицата е резулатът от работата на алгоритъма:

aa	amam andand	eggegg	ifif	thethe twotwo
.22 a	.18 am .20 and	. 05 egg	025 if	.02 the .08 two
a.am	amand andegg	eggif	ifthe	thetwo
.58 a	.56 and .30 and	.35 if	.29 if	.12 two

a and	amegg and.	eggthe	iftwo
1.02 am	.66 and .80 is	f .39 if	.47 if

a..egg am.. if and..the egg..two
1.17am 1.21 and .84 if 57 if

цена

a.. if am.. the and..two 1.83 and 1.27 and 1.02 if

Корен на оптималното (до момента) бинарно д-во

a.. the am..two 1.89 and 1.53 and

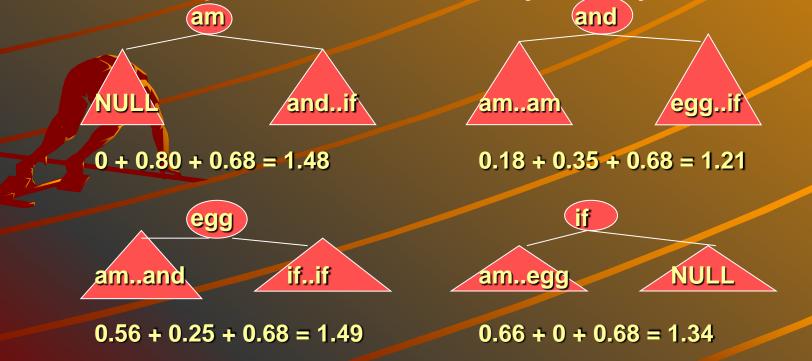
Определянето на цената на това поддърво (за подредица от am до if) е показано на следващ слайд

a.. two 2.15 and Цена / корен на оптимално д-во, съдържащо всички думи (третият граф)

за всяка подредица от думи се пази и цена (в случая вероятност) и корен на оптималното (до момента) binary search tree.

Най-долу се получава цена и корен за оптималното дърво, съдържащо всички думи (това е и оптималното дърво, показано трето в графите).

Как са изчислени стойностите за всички възможни корени до откриване на оптимално дърво за подмножество от думи между: am .. if :



Последователно се поставят за корени думите: am, and, egg, if и се изчислява цена за дървото - оттам откриваме минималната цена.

Например, когато and е корен, лявото поддърво ат...ат има цена 0,18 (вече определена), дясното – egg-if с цена0.35 и

 $p_{am}+p_{and}+p_{egg}+p_{if}=0.68$ (винаги)

Тогава общата цена е 1,21

3

Времето е O(N) защото имаме тройно вложен цикъл



4.3* Най-къс път м/ду всички двойки точки в граф

Имаме насочен граф G=(V.E).

Алгоритъмът на Dijkstra намиращ най-къси пътища от произволен връх до всички останали (O(IVI_{^2}) работеше на етапи: започваме от връх s; всеки връх (v) се селектира като междинен. За всеки връх w-дестинация, дистанцията d се изчислява така:

dw= min (dw,dv+cvw)

Сега селектираме върхове последователно. Нека D к,і,ј е теглото на найкъсия път от vi до vj през v1,v2...vk (к междинни върха).

T.e.: Do,i,j = Ci,j или безкрайност ако няма ребро.

DIVI,i,j е най-къс път от vi до vj из целия граф.

Когато к>0, можем да напишем оптимизирана програма за най-къс път

D k,i,j.

Най-късият път е или този който изобщо не минава през v1,v2...vk междинно, или получен от сливане на 2 пътя: vi →vk; vk→vj, всеки минаващ през първите k-1 междинни върхове.

```
for (int k=0;k<n; k++)
        // нека всеки връх разгледаме като междинен
         for (int l=0; l < n; l++)
                 for(int j=0; j < n; j++)
                          if(d[i][k] +d[k][j] < d[i][j];
                          // нов най-къс път
                           d[i][j] = d[i][k] + d[k][j];
                           path[ i ][ j ] = k;
или формулата:
D_{k,l,j} = min\{D_{k-1,i,j}, D_{k-1,i,k} + D_{k-1,k,j}\}
Времето отново e O( V )
```

Освен това алгоритъмтът е отличен за разпаралеляване.

5. Алгоритми с backtracking

Достига до добри решения, но е непредвидимо бавен, поради връщанията

Пример: подреждане на мебели в къща.

спира се на различни етапи и се кара до изчерпване

възможностите.

4.1пример: проблем – реконстркция (приложение във Физика, молекулярна биология..)

Нека имам N точки: p1,p2...pn подредени по оста x. Дадени разстоянията : на брой N(N – 1) / 2. Да приемем: x1 = 0 Ако имахме дадени координатите, лесно можем да изчислим

разстоянията. Оценката е O(N)

Нека дистанциите с дадени сортирано.

Задачата е: да се реконструират координатите на точките от дистанциите между тях.

дадени:
$$D = \{1,2,2,2,3,3,3,4,5,5,5,6,7,8,10\}$$
 D има 15 елемента , следователно, $N=6$

$$x1 = 0$$

 $D = \{1,2,2,2,3,3,3,4,5,5,5,6,7,8\}$

най-голямата оставаща дистанция е 8, следователно, или x2=2 или x5=8. Откъдето и да тръгнем, все ще стигнем края – ако съществува (връщанията от неуспех са равновероятни).

Нека
$$x5=8$$
. $x5=x1=8$

$$x1 = 0$$
 $x5 = 8$ $x6 = 10$ $D = \{1,2,2,3,3,3,4,5,5,5,6,7\}$

Сега 7 е най-голямото оставащо разстояние. Или х4=7 или х2=3 (за да има разст. = 7 от х2 до х6). Ако х4 = 7 – проверяваме разст. до х6 и х5 и виждаме че ги има във вектора с разст. Ако решим х2=3, то х3-х1=3; х5-3=5 - също са във вектора. Значи и това става.

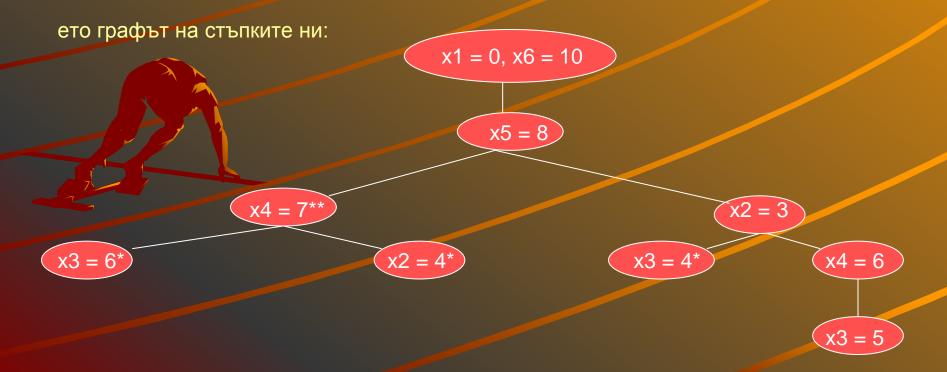
Избираме едното за да продължим. Нека х4=7

Сега макс. разст. е 6, така че или х3=6 или х2=4. Проверяваме х3=6 - не може. Проверяваме х2=4 – не може. Връщаме се назад. х4=7 отпадна. Сега опитваме х2=3

сега остава да изберем или х4=6 или х3=4. х3=4 е невъзможно. Значи остава х4=6:

$$x1 = 0$$
 $x2 = 3$ $x4 = 6$ $x5 = 8$ $x6 = 10$ $D = \{1,2,3,5,5\}$

достигнахме успешен край! Векторът от разстояния се изпразни



Забележки:

избраното решение води до разстояния, които липсват в D.
 върхът има само неуспяващи деца, следователно този път се изоставя.

Поради backtracking анализът е вероятностен от... до...

оценка:

Ако няма връщане O(N logN)

5.2 Елементи от теория на игрите

1.изчерпателен анализ на ходовете – възможно ли е? Примери: компютърен шах, игра: крави /бикове крави/бикове е лесна за програмиране, защото броят е изчерпаем, известни са слабите ходове (могат да се вкарат в таблица). Алгоритъмът никога не греши (никога не губи) и винаги ще спечели ако му се даде възможност.

Стратегия МИНИ/Макс

разглеждаме "силата" на позициите: Ако води победа – оценката е +1, ако равенство: 0; ако е позиция с която компютърът губи: −1.
Такива позиции са терминални.
За останалите → рекурсивно се играе по възможните ходове.

Значи стратегияте е : противникът се стреми да минимизира стойността (някаква оценъчна функция), Играчът (компютърът) – да я максимизира. Пробват се всички възможни в момента ходове. Избира се този с макс. стойност. За ход на противника – същото: прохождат се всички ходове и се избира този с минимизираща стойност.

```
1/2
int TicTakToe:: findCompMove( int & bestMove)
        { int i, response Value;
                                                                      Компютърна
        int dc /// don't care: стойността не се използва
                                                                      селекция
        int value;
        if(fullBoard())
                 value = DRAW;
                                             //равенство
        else
        if( immediateCompWin( bestMove )) //има възможност за редица
                 return COMP_WIN
        else
                 { value = COMP_LOSS; bestMove = 1;
                 for( i = 1; i < 9; i++) // пробва се с всяко квадратче
                          if( isEmpty( i ))
                                   place(i, COMP);
                                   responseValue = findHumanMove( dc );
                                   unplace(i);
                                   if(responseValue > value) // търси подобрение
                                            //update best move
                                    value = responseValue;
                                   bestMove = i;
        return value;
```

```
ето играта на човека (двете функции рекурсивно се викат, като оценките им за
успешен ход са противоположни):
int TicTacToe::findHumanMove( int & bestMove)
                                                                        2/2
        int i, responseValue;
        int value; int dc //don't care
        if( fallBoard())
                 value = DRAW;
         else
                 if( immediateHumanWin( bestMove) )
                          return COMP_LOSS;
                 else
                          { value = COMP_WIN; bestMove = 1;
                          for( i = 1; i <= 9; i++ ) //изпробва всяко квадратче
                                  if( isEmpty( i ) )
                                   place( i, HUMAN);
                                   responseValue = findCompMove( dc );
                                   unplace(i); //restore board
                          ii( response Value < value ) // търси намаляване
                                           // update best move
                                   value = responseValue
                                   bestMove = i;
                          }} return value;
```

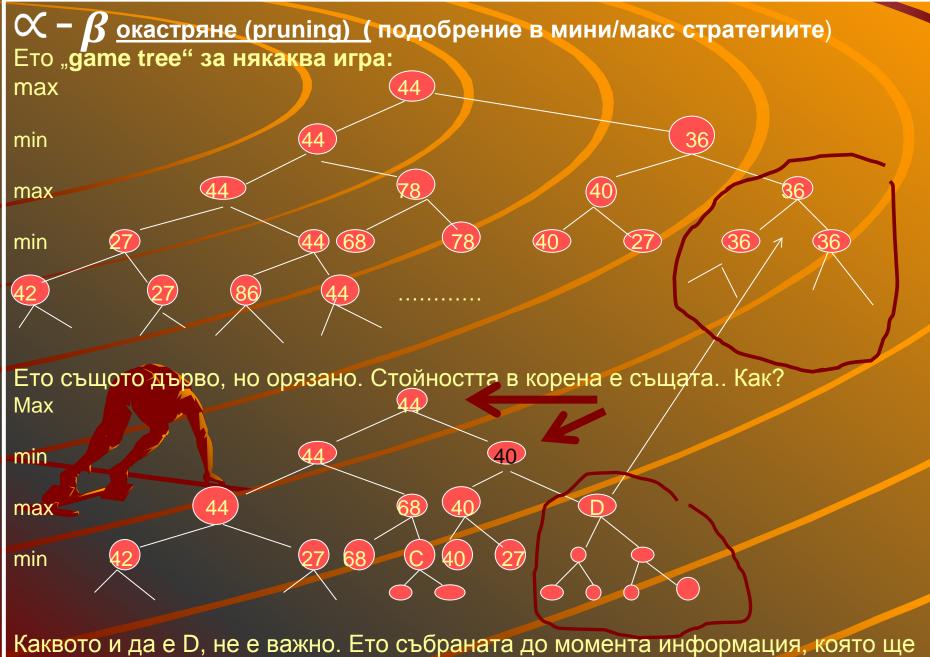
Най- много изчисления са нужни когато компютърът започва играта (много възможности за проверка). Избира се позиция 1 (горе, ляво, макар че това е условно). Има 97,162 възможни продължения за анализ.

Ако компютърът играе втори – позициите за анализ са намалели на 5185. Ако е бил избран центърът: 9761, при избор на ъгъл: 13,233.

Очевидно при шах тази стратегия за прохождане (до терминален) е непосилна (повече от $10^{100}\,$ позиции за анализ). Спира се донякъде, вика се функция, оценяваща достигнатата позиция. (напр. проходимост на фигури, качество на фигури и т.н.) . Тази функция е основната за шах- програма.

Много е важно колко хода напред могат да се прегледат (дълбочината на рекурсията). За да се увеличат:

в таблица се пазят вече анализирани (и аналогични) ходове. Когато се срещнат – се скача към терминален.



Каквото и да е D, не е важно. Ето събраната до момента информация, която ще служи за изчисляване на D:



Стратегията е лесна за реализация и силна. Колкото по-нагоре в дървото я приложим, толкова ефектът е по-добър.

Ограничава търсенето до О(√N) възли, където N е размера на цялото дърво. Това позволява удвояване броя на прохожданите ходове. Играта крави/бикове не е подходящ пример за ползата (има много идентични (като оценка) ходове). Въпреки това прилагането й в началния момент свежда от 97162 до 4493 възлите за разглеждане (само нетерминални възли се разглеждат при тази стратегия).

 $_{\rm g} \propto f$

```
Int TicTacToe::findCompMove(int &bestMove, int alpha, int beta)
         { int i, responseValue;
         int dc
                                                                   Преди повикване:
                                                                   alpha = COMP_LOSS
         int value;
                                                                   beta = COMP_WIN
         if( fullBoard() )
                          value = DRAW;
         else
                  if( immediateCompWin( bestMove)
                  return COMP_WIN; //bestMove се устан. om immediateCompWin()
                  else
                            value = alpha; bestMove = 1;
                            for( i = 1; i <= 9 && value < beta; i++)
                            { if( isEmpty( i ) )
                                               place ( i, COMP );
                                         responseValue = findHumanMove( dc );
                                               unplace( i ); //restore board
                                               if( responseValue > value )
                                               { // update best move
                                               value = responseValue;
                                               bestMove = i;
         return value;
```