OS202 - Systèmes parallèles et distribués

Examen de Rattrapage

ODORISSIO PEREIRA, Vitor

2024

Computer Specifications

- CPU: AMD Ryzen 7 6800H

    – 8 Physical Cores

    – 16 Logic Cores

    – L1 Cache: 64KB (32K data + 32K instructions) / core

    – L2 Cache: 512KB / core

    – L3 Cache: 16MB shared

The physical cores of a CPU are the real components inside it that can perform logical operations, while the logical cores represent the number of threads a CPU can handle "simultaneously". By assigning multiple threads to a single physical core, the CPU can utilize its idle time more efficiently.

Classificador de Bayes

IV.1

The algorithmic complexity for calculating the joint distribution is:

$$O\left(K_1 \times \prod_{i=1}^{n} n_i \times 3 \right)$$

The storage complexity for storing the joint distribution is:

$$O\left(\prod_{i=1}^{n} n_i \times 3 \right)$$

IV.3

The parallelization did not result in significant speedup; in fact, it was slower than the sequential code, as seen in illustration 1. This is likely due to the communication overhead outweighing the computational benefits of parallelization. Using shared memory between the threads could help in this particular case.
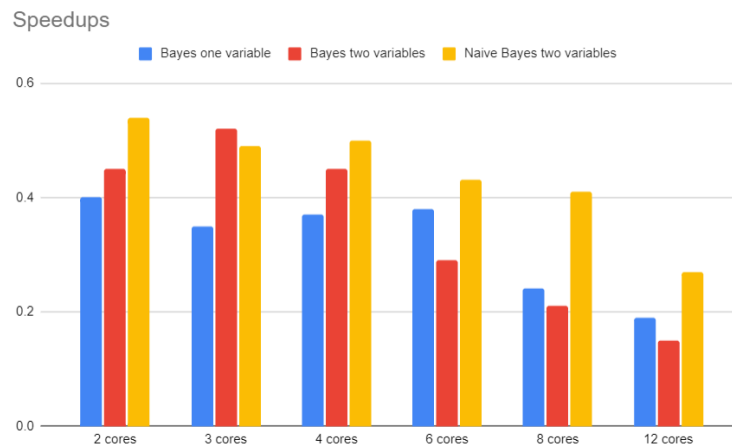
Illustration 1 – Speedup Values

Output with 8 process:

Bayes one variable jour:

Sequential Précision de la prédiction jour: 44.21658121795674%

Sequential Temps bayes une variable : 0.11952042579650879 secondes

Parallel Précision de la prédiction jour: 44.21658121795674%

Parallel Temps bayes une variable : 0.6326315402984619 secondes

Speedup for Bayes one variable: 0.19

Bayes one variable mois:

Sequential Précision de la prédiction mois: 57.35067316910508%

Sequential Temps bayes une variable : 0.12831449508666992 secondes

Parallel Précision de la prédiction mois: 57.35067316910508%

Parallel Temps bayes une variable : 0.5076053142547607 secondes

Speedup for Bayes one variable: 0.25

Bayes two variables jour/mois:

Sequential Précision de la prédiction jour/mois: 61.12292109541078%

Sequential Temps bayes deux variables : 0.2805595397949219 secondes

Parallel Précision de la prédiction jour/mois: 61.12292109541078%

Parallel Temps bayes deux variables : 1.3371813297271729 secondes

Speedup for Bayes two variables: 0.21

```
Sequential Précision de la prédiction jour/mois (naïf): 59.51815264057355%
Sequential Temps bayes naïf deux variables : 0.19003891944885254 secondes
Parallel Précision de la prédiction jour/mois (naïf): 59.51815264057355%
Parallel Temps bayes naïf deux variables : 0.5360288619995117 secondes
```

V.1

To parallelize the decision tree construction, the best way is to split the work when the tree creates child nodes. After the tree splits at one node, the left and right child nodes can be built separately and at the same time. This allows the program to work on different parts of the tree in parallel.

The problem with parallelizing this algorithm is that the deeper you go in the tree, the less data there is to work with. You can parallelize, there is just not enough work in the deeper nodes to make a big difference, limiting the algorithim. Also, the top of the tree still needs to be built in sequence, so this limits how much faster the parallelization can be.

VI.1

```
Ratio des différents modes de production électrique :
Carbonne 36.83640303358613%
Decarbonne 29.419118259854987%
Normale 33.74447870655888%


Temps de construction de l'arbre de décision : 17.718921661376953 secondes


Temps pris pour calculer les predictions : 0.047512054443359375 secondes


Précision de l'arbre de précision': 72.22708515693385%
```

RÉFÉRENCES