Experimento velocidad de ejecución

**1. Proceso o sistema que se estudia y variables respuestas de interés**

El proceso estudiado es la velocidad de ejecución de código escrito en C usando y no usando OpenMP, adicionalmente se compararán estas salidas con ejecución de código escrito en Python y R.

Definiciones básicas: OpenMP es una interfaz de programación de aplicaciones para la programación multiproceso de memoria compartida en múltiples plataformas.

Lenguajes de programación usados: C, Python (Numpy) R.

Las variables que intervienen en este experimento son:

* Computadora usada.
* Lenguaje usado.
* Algoritmo usado.

La variable respuesta corresponde al tiempo medido en segundos de una multiplicación de matrices NxN con valores entre 1 y 100 generadas aleatoriamente, estas ejecuciones se desarrollaron una vez reiniciada la computadora, desconectando el internet y cualquier evento, esperando a que todos los procesadores (CPU) estuvieran aproximadamente al 100% de sleeping.

**2. Objetivo del experimento**

Comparar y evidenciar existencia de reducción de tiempo usando OpenMP. Se espera que el experimento responda a las siguientes preguntas.

* ¿Existe diferencia en tiempo en la ejecución de algoritmo escrito en C con OpenMP y sin OpenMP?
* ¿Existe diferencia en tiempo en la ejecución de los algoritmos escritos en C y los algoritmos escritos en Python (Numpy) y R.

No es posible generalizar resultados a otros tipos de lenguajes a los usados; pues cada lenguaje puede diferir en su forma de ejecución. Tambien debido a las limitaciones de equipos, no se puede esperar que las condiciones de experimentación emulen las condiciones úsales a las que están ejecutando dichos algoritmos. Las conclusiones que se obtengan sólo pueden ser discutidas en términos de las condiciones puestas en este experimento, aunque ellas podrían dar indicios de los resultados que se darían bajo condiciones más normales.

**3. Identificación de todas las fuentes de variación:**

**3.1. Factores de estudio y sus niveles**

El factor de tratamiento “tipo de algoritmo” se han elegido tres lenguajes de programación: C, C con OpenMP, Python (Numpy) y R. Note que el experimentador no controla la existencia de deamons o tareas sobre el pc; asume que los pc usados serán típicos de la población disponibles en las tiendas. Otro factor de tratamiento “tamaño de matriz” se han elegido 5 tamaños de matriz: 100x100, 200x200, 300x300, 400x400 y 500x500, con valores aleatorios entre 1 y 100. Si esta presunción no es cierta, entonces los resultados del experimento no tendrán una aplicación general. Se tomarán 30 veces el tiempo por algoritmo con el fin de que el experimento sea tan representativo como pueda ser posible de la población de ejecuciones.

**3.2. Factores de bloque, de ruido y covariables**

A parte de las diferencias en la forma de ejecución del algoritmo en los diferentes lenguajes, los procesadores pueden variar, adicionalmente el calor generado por el pc, y la temperatura del día no fueron necesariamente todos expuestos a la misma cantidad de calor. Sin embargo, el experimentador no sintió que las unidades experimentales serían suficientemente variables como para justificar bloqueo. Estos pudieron haber sido usados como covariables, pero el experimentador eligió en cambio medir los cambios de tiempo, es decir, el tiempo final menos el tiempo inicial.

Otras fuentes de variación incluyen generación de números aleatorios, imprimir en pantalla. Todas estas fueron consideradas menores. Ningún factor de ruido fue incorporado al experimento.

**3.3. Unidades experimentales**

El experimento se llevará a cabo usando diferentes pc, a los que se desconecta las redes, bluetooth y cualquier otro tipo de conexión inalámbrica, los cuales se reinicia el pc, al verificar que los CPU’s estén aproximadamente al 100% sleeping se ejecutan los algoritmos, estos se ejecutan por consola. Se hará un registro de los tiempos de ejecución en segundos.

Un estudio piloto indicó que esta manera de ejecución es suficiente para cubrir cualquier variación que el pc podría dar a los tiempos de ejecución de los algoritmos por los tamaños de algoritmo.

**4. Regla de asignación de las unidades experimentales a los tratamientos**

Un número igual de observaciones será hecha sobre cada uno de los niveles de los factores. Por tanto, se prepararán ***n*** ejecuciones por cada algoritmo y tamaño de matriz. Este tamaño elegido fue de 30 para cumplir con el **TLC**. No se aleatorizarán por facilidad de ejecución.

**5. Procedimiento experimental, mediciones y dificultades anticipadas**

Los algoritmos son compartidos por GitHub para la reproducibilidad del experimento, al verificar que los CPU’s estén aproximadamente al 100% sleeping se ejecutan los algoritmos y desconectar las redes, bluetooth y cualquier otro tipo de conexión inalámbrica

Después que se han ejecutado los algoritmos, en pantalla se muestra el tiempo de ejecución en segundos. Estos tiempos serán registrados. El análisis se realizará sobre las diferencias en el uso de OpenMP.

**Dificultades experimentales esperadas:**

* Programas, tareas o deamons en el pc. Por tanto, los datos pueden no mostrar diferencias en los tiempos.
* Tiempos afectados por la ejecución anterior, correlación entre muestras.
* Tiempos mayores a los esperados en el estudio piloto.

**Dificultades encontradas durante la experimentación:**

* En el estudio piloto se evidencio tareas que no fueron apagadas y afectaban el tiempo como el internet o el bluetooth de los audífonos.
* El uso de consola disminuía los tiempos, usar herramientas IDE aumentaba tiempos.
* En el estudio piloto se evidencio correlaciones entre los tiempos.

**6 Experimento piloto**

Se corrió un experimento piloto con dos propósitos: Primero, para identificar las dificultades listadas en la Sección 5, segundo, para obtener una estimación de la varianza para estimar el número de réplicas. La varianza del error fue estimada aproximadamente de 0.146 s2, el valor del MSE en dicho experimento piloto. De hecho, este es una sobre estimación y hubiese sido mejor usar un intervalo de confianza unilateral para la varianza.

**7. Especificación del modelo estadístico, análisis o pruebas a realizar y tamaños de muestra**

Dado que se tendrá cuidado en controlar todas las fuentes extrañas de variación, se asumió que el siguiente modelo es una aproximación razonable:

 (1)

Nota: Me equivoque y es i=j=2

Con el fin de responder a la pregunta sobre las diferencias en los pesos, un ANOVA de un factor de efectos fijos será calculada al 0.05 de significancia para probar:

 (2)

Nota: Me equivoque y es 1 y 2 el 3 no

Donde el estadístico de prueba y su distribución es normal y el criterio de decisión con valor P es 0.05 Si se detecta significancia estadística, se estimarán las medias y los efectos según tipo algoritmo, así como sus intervalos de confianza del 95%. Para hallar mejor las diferencias entre pares de tratamientos, se construirán los I.C de Tukey simultáneos del 95%.

**Cálculo del número de observaciones necesarias:** Se halla ***n=30*** para cada algoritmo y tamaño de matriz (se consideró importante detectar una diferencia en disminución de tiempos de al menos xxx% entre los algoritmos en C con OpenMP y sin OpenMP., con una probabilidad de 0.90 y nivel de significancia de 0.05).

**Revisión de las decisiones anteriores:** No es difícil obtener treinta observaciones de cada uno de los tratamientos; pequeños ajustes al procedimiento experimental que fueron considerados como necesarios de realizar durante el experimento piloto ya han sido incorporados.

**8. Datos recolectados y análisis estadísticos de resultados**

A continuación, se presentan los datos en el orden de recolección (el molde asignado) según procedimiento de aleatorización indicado en la Tabla 1 y la tabla resumen de los datos recolectados durante el experimento.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Tabla 1. Procedimiento de aleatorización usado para asignar los moldes a las réplicas | | | | | |
| Asignación de número aleatorio a réplicas a realizar | | Asignación molde a réplicas según orden de número aleatorio asignado | | |  |
| Tratamiento | Número aleatorio | Tratamiento | Número aleatorio | Molde | Observaciones |
| 1 | 0,4435 | 3 | 0,0265 | 1 | 1,86 |
| 1 | 0,3069 | 1 | 0,3069 | 2 | -0,30 |
| 1 | 0,7066 | 3 | 0,3440 | 3 | 2,03 |
| 1 | 0,6581 | 1 | 0,4435 | 4 | -0,10 |
| 2 | 0,5490 | 2 | 0,5201 | 5 | 2,63 |
| 2 | 0,5201 | 2 | 0,5490 | 6 | 2,61 |
| 2 | 0,8509 | 2 | 0,6344 | 7 | 2,41 |
| 2 | 0,6344 | 1 | 0,6581 | 8 | -0,14 |
| 3 | 0,0265 | 1 | 0,7066 | 9 | 0,40 |
| 3 | 0,8009 | 3 | 0,8009 | 10 | 2,26 |
| 3 | 0,3440 | 2 | 0,8509 | 11 | 3,12 |
| 3 | 0,9545 | 3 | 0,9545 | 12 | 1,82 |

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Tabla 2. resumen de observaciones por tratamiento | | | | | |
| Jabón | Pérdida de peso (grs.) | | | | Promedio |
| 1  (regular) | -0.30 | -0.10 | -0.14 | 0.40 | -0.0350 |
| 2  (desodorante) | 2.63 | 2.61 | 2.41 | 3.15 | 2.7000 |
| 3  (hidratante) | 1.86 | 2.03 | 2.26 | 1.82 | 1.9925 |

Vea a continuación el análisis de estos datos.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Tabla 3. ANOVA | | | | | |
| Fuente | DF | Sum Sq | Mean Sq | F value |  |
| jabón | 2 | 16.12 | 8.06 | 104.45 | 5.914×10-7 |
| Residuals | 9 | 0.69 | 0.08 |  |  |

|  |  |
| --- | --- |
| Figura 1. Distribución de la respuesta según tipo de jabón | Figura 2. Representación de los I.C de las diferencias de medias  del 95%, tipo Tukey |

De la Figura 1 es claro que puede existir una diferencia significativa entre las medías del peso perdido por los jabones regulares, desodorante e hidrantes, lo cual es corroborado por los resultados del Test ANOVA en la Tabla 3 ya que…..

De acuerdo con la Figura 2, también se concluye por los I.C de Tukey del 95% que hay diferencias significativas entre las medias de pérdida de peso de los tres tipos de jabones (ver también Tabla 6) puesto que….

En la Tabla 4 se presentan las estimaciones para las medias de pérdida de peso para cada tipo de jabón mientras que la Tabla 5 presenta las estimaciones de los efectos individuales con sus I.C del 95% y resultados para el test de significancia individual de estos efectos. Se observa que para el jabón tipo 1 la pérdida media es negativa así como el efecto estimado (es decir, ganó peso), esto ocurre debido a que….

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Tabla 4. Medias estimadas e I.C del 95% | | | | | |
| jabon | lsmean | SE | df | lower.CL | upper.CL |
| 1 | -0.0350 | 0.1389 | 9 | -0.3492 | 0.2792 |
| 2 | 2.7000 | 0.1389 | 9 | 2.3858 | 3.0142 |
| 3 | 1.9925 | 0.1389 | 9 | 1.6783 | 2.3067 |

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| Tabla 5. Efectos de tratamientos estimados e I.C del 95% | | | | | | |
| Tipo jabón | Estimate | Std. Error | t value |  | lower CI | upper CI |
| jabon:efecto regular | -1.59 | 0.11 | -14.00 | 2.051×10-7 | -1.84 | -1.33 |
| jabon:efecto desodorante | 1.15 | 0.11 | 10.12 | 3.245×10-6 | 0.89 | 1.40 |
| jabon:efecto hidratante | 0.44 | 0.11 | 3.88 | 3.734×10-3 | 0.18 | 0.70 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Tabla 6. Comparaciones múltiples de medias método Tukey con I.C del 95% | | | | |
| Comparación | diff | lwr | upr | Pvalue-adj |
| 2-1 | 2.7350 | 2.186547 | 3.2834532 | 0.0000006 |
| 3-1 | 2.0275 | 1.479047 | 2.5759532 | 0.0000074 |
| 3-2 | -0.7075 | -1.255953 | -0.1590468 | 0.0142876 |

Para la validación de los supuestos de normalidad sobre los errores se presentan los resultados del test Shapiro Wilk (ver Tabla 6) y el gráfico de probabilidad normal con residuos de ajuste estandarizados (ver Figura 3). Para el supuesto de varianza constante , con , se presentan los gráficos de residuales estandarizados en la Figura 3 y los resultados del test de homogeneidad de Levene exhibidos en la Tabla 7, usando las desviaciones absolutas con relación a la mediana de cada nivel. Con base en estos resultados se concluye que….



Figura 3. Gráficos de residuales (arriba) y de probabilidad normal

|  |
| --- |
| Tabla 7. Resultados Test Shapiro Wilk y de Levene |
| Test Shapiro Wilk con residuos estandarizados  W = 0.85898, p-value = 0.04747 |
| Test de Levene: Desviaciones absolutas de la mediana   |  |  |  |  | | --- | --- | --- | --- | | Fuente | Df | F value |  | | group | 2 | 0.04 | 0.9620 | | Error | 9 |  |  | |

**9. Conclusiones y recomendaciones**

Con base en los resultados experimentales, los principales hallazgos fueron….

Como recomendación para mejorar este experimento puede considerarse lo siguiente…

**10. Programa R**