Tabela 1: Configurações testadas.

Configuração	Duração tabu	Solução inicial	Compilação
1	$0.1n$	Aleatória	-O
2	$0.1n$	Aleatória	-O3
3	$0.2n$	Aleatória	M_1
4	M_2	Gulosa	M_1

Soluções 3

O plano de teste deve incluir diversas instâncias. Como o método GSAT/Tabu é randomizada (por selecionar um dos melhores vizinhos aleatoriamente) vamos executar todas configurações múltiplos vezes e comparar os tempos médios. Isso permitiria também usar uma busca tabu com duração tabu aleatória, mas para reduzir o número de testes vamos usar uma duração tabu de $0.1n$ e $0.2n$. Concretamente o plano de teste é

Instâncias Classes de instâncias uf75, flat75, RTI_k3_n100_m429, e instânicas 20, 40, 60, 80, 100 de cada classe.

Replicações 30 com sementes $1, \ldots, 30$ e tempo limite 120s por replicação.

Configurações Testaremos as configurações da tabela 1, com M_1 o melhor entre -O e -O3 conforme o resultado do primeiro teste e M_2 o melhor entre $0.1n$ e $0.2n$ conforme o resultado do segundo teste.

A tabela 2 mostra os resultados de todos testes.

Teste 1 e 2 O primeiro teste compara as colunas 1 e 2 da tabela 2. A otimização completa é (um pouco) mais rápida em 10 dos 15 testes, que não é suficiente para rejeitar a hipótese nula que eles precisam o mesmo tempo usando um teste de sinal ($p = 0.15$), sendo a hipótese alternativa que a otimização completa precisa menos tempo. Como esperado o número de passos foi idêntico em todas instâncias resolvidas (882 do 900 testes), logo não podemos concluir rejeitar a hipótese nula que ambas níveis de otimização precisam o mesmo número de passos. Nos vamos manter a otimização completa para os próximos testes.

Teste 3 A tabela 2 mostra os resultados do terceiro teste (colunas com configurações 2 e 3, com $M_1 =$ -O3). A configuração 2 é mais rápida em 8 dos 15 casos, e novamente não podemos rejeitar a hipótese nula que eles precisam o mesmo tempo (isso também com um teste de postos com sinais), mesmo com médias diferentes. Vamos escolher uma duração tabu $0.2n$ para o último teste.

Teste 4 A tabela 2 mostra os resultados do quarto teste (colunas com configurações 3 e 4, com $M_2 = 0.2n$). A configuração 4 é mais rápida em 8 dos 15 casos, e novamente não podemos rejeitar a hipótese nula que eles precisam o mesmo tempo (isso também com um teste de postos com sinais).

Tabela 2: Comparação otimização simples (-O) e otimização completa (-O3).

Nome	Inst.	Tempo	Tempo	Tempo	Tempo
Configuração		1	2	3	4
RTI_k3_n100_m429	20	18.93	18.53	18.67	30.37
RTI_k3_n100_m429	40	4104.20	4101.83	13.40	17.70
RTI_k3_n100_m429	60	8.90	8.57	18.07	16.87
RTI_k3_n100_m429	80	18.67	18.27	8.93	7.17
RTI_k3_n100_m429	100	13.50	13.23	15.33	11.27
flat75	20	13.87	13.87	318.17	424.83
flat75	40	129.73	129.73	3394.07	3690.30
flat75	60	10.60	10.63	263.70	314.63
flat75	80	46.47	46.60	1209.10	832.33
flat75	100	17.70	17.90	315.23	487.50
uf75	20	24010.37	24010.33	2.70	2.53
uf75	40	5509.67	5474.10	24.13	14.70
uf75	60	19.40	19.03	3.10	1.73
uf75	80	6.03	5.97	1.13	1.23
uf75	100	4258.80	4253.13	5.10	4.57
Médias		2545.79	2542.78	374.06	390.52

Complexidade empírica A complexidade empírica do configuração 3 no modelo linear é

$$1.22\mu\text{s} \cdot n^{6.63} m^{-2.35}.$$

Para testar a linearidade em m a hipótese nula é $\beta = 1$ (com $\hat{\beta} = -2.35$) e a hipótese alternativa $\beta \neq 1$. Um teste t (de acordo com o exemplo 6.14 das notas) rejeita a hipótese nula num nível de $p < 0.001$, usando

```
2*pt(t,n-2,lower.t=F)
```

para o teste bicaudal.
(O modelo linear a princípio é suspeito, por se tratar de um problema NP-completo. Logo a hipótese exponencial seria mais adequada. Neste caso obtemos o modelo

$$70\text{ms} \cdot 0.96^n 1.02^m.$$

Ambos os modelos não são muito confiáveis por serem baseados em uma amostra pequena demais.)