Application of Nested Monte-Carlo methods to the Traveling Salesman Problem with Time Windows

Tristan Cazenave\textsuperscript{1} and Fabien Teytaud\textsuperscript{1,2}

\textsuperscript{1} LAMSADE, Université Paris Dauphine
\textsuperscript{2} HEC Paris, CNRS, 1 rue de la Libération 78351 Jouy-en-Josas
Outline

- Traveling Salesman Problem with Time Windows
- Nested Monte-Carlo Algorithm
- Nested Roll-out Policy Adaptation
- Experiments
- Conclusion
Traveling Salesman Problem (TSP)

- **Data**
  - List of cities
  - Distances between all cities

- **Goal**
  - Find a path visiting each city exactly once
  - The path must be as short as possible
Traveling Salesman Problem with Time Windows (TSPTW)

- Additionnal property: Time windows
  - A city can not be visited before a certain time and after a certain time

- Some problems have no solution

- Finding a valid solution is NP-hard
Outline

- Traveling Salesman Problem with Time Windows
- Nested Monte-Carlo Algorithm
- Nested Rollout Policy Adaptation
- Experiments
- Conclusion
Nested Monte-Carlo (NMC)

[Cazenave, 2009]

- Tree exploration algorithm
- Evaluation with Monte-Carlo simulations
- Particularly efficient for one player games and when late decisions are as important as early ones.
Nested Monte-Carlo (NMC)

[Cazenave, 2009]

- Nested plays a whole game and returns the associated score
- Nested takes for parameters the level $n$ and the current position (recursive algorithm)
- Principle
  - The score of an action is calculated by calling a nested with level $n-1$
  - The level 0 of NMC is a Monte Carlo simulation (random play until the end of the game)
NMC

- **Level 0**
  - Monte-Carlo policy
  - Choose a city randomly

- **Level > 0**
  - Launch NMC($level-1$)
  - The action with the highest score is chosen
NMC(level=1) example
Adding Heuristics

[Rimmel et al, 2011]

- The algorithm can be improved by modifying the Monte Carlo simulations.
- Instead of uniformly random, the actions are chosen according to expert knowledge:
  - The distance to the last city
  - The waiting time (related to the inf bound of the time window)
  - The remaining time before the end of the time window
Outline

- Traveling Salesman Problem with Time Windows
- Nested Monte-Carlo Algorithm
- Nested Rollout Policy Adaptation
- Experiments
- Conclusion
Nested Rollout Policy Adaptation (NRPA)

- NMC can be improved by modifying the Monte Carlo simulations.

- Instead of random playouts, a policy is learned:
  - Increase the weights of the best cities
  - Decrease the weights of other cities
  - For each city: compute a probability proportional to the exp of its weight

[Rosin, 2011]
Nested Rollout Policy Adaptation (NRPA)

- **Level 0**
  - Adapted policy
  - Choose a city accordingly to its probability

- **Level > 0**
  - Do N iterations of NRPA(level -1)
  - Update
    - The scores
    - The sequences
    - The policy
Adding expert-knowledge (NRPA_EK)

- Force to visit cities as soon as they go after their windows end.
- Avoid visiting a city if it makes another city go after its windows end.
- Consider all moves if no move available after these two tests.
- Important point: Optimal moves can not be pruned with this expert knowledge.
Outline

- Traveling Salesman Problem with Time Windows
- Nested Monte-Carlo Algorithm
- Nested Rollout Policy Adaptation
- Experiments
- Conclusion
Experiment protocol

- Experiments:
  - Tuning of NMC
  - Analyzes of N and the level (NRPA)
  - Comparison of NRPA and NRPA_EK on one problem.
  - Comparison of the best results found by NMC, NRPA and NRPA_EK on a set of standardized problems

[Lopez-Ibanez and Blum, 2010]
Experiments (Tuning of NMC)

<table>
<thead>
<tr>
<th>Iterations</th>
<th>BEST</th>
<th>KBEST</th>
<th>MEAN</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2.7574e+06</td>
<td>2.4007e+06</td>
<td>2.3674e+06</td>
</tr>
<tr>
<td>2</td>
<td>5.7322e+04</td>
<td>3.8398e+05</td>
<td>1.9397e+05</td>
</tr>
<tr>
<td>3</td>
<td>7.2796e004</td>
<td>1.6397e+05</td>
<td>618.22</td>
</tr>
<tr>
<td>4</td>
<td>5.7274e+04</td>
<td>612.60</td>
<td>606.68</td>
</tr>
<tr>
<td>5</td>
<td>2.4393e+05</td>
<td>601.15</td>
<td>604.10</td>
</tr>
<tr>
<td>6</td>
<td>598.76</td>
<td>596.02</td>
<td>602.96</td>
</tr>
<tr>
<td>7</td>
<td>599.65</td>
<td>596.19</td>
<td>603.69</td>
</tr>
<tr>
<td>8</td>
<td>598.26</td>
<td>594.81</td>
<td>600.79</td>
</tr>
<tr>
<td>9</td>
<td>596.98</td>
<td>591.64</td>
<td>602.54</td>
</tr>
<tr>
<td>10</td>
<td>595.13</td>
<td>590.30</td>
<td>600.14</td>
</tr>
<tr>
<td>11</td>
<td>590.62</td>
<td>591.38</td>
<td>600.68</td>
</tr>
<tr>
<td>12</td>
<td>593.43</td>
<td>589.87</td>
<td>599.63</td>
</tr>
<tr>
<td>13</td>
<td>594.88</td>
<td>590.47</td>
<td>599.24</td>
</tr>
<tr>
<td>14</td>
<td>590.60</td>
<td>589.54</td>
<td>597.58</td>
</tr>
<tr>
<td>15</td>
<td>589.07</td>
<td>590.07</td>
<td>599.73</td>
</tr>
</tbody>
</table>

Table 1. Evolution of the true score on the problem rc206.3.
Experiment results (1)
Experiment results (2)

- Hardest problem from the set,
- 46 cities,
- Best known result: 868,76
## Experiment results (3)

<table>
<thead>
<tr>
<th>Problem</th>
<th># cities</th>
<th>State of the art</th>
<th>NMC_EK score</th>
<th>NRPA score</th>
<th>NRPA_EK score</th>
</tr>
</thead>
<tbody>
<tr>
<td>rc206.1</td>
<td>4</td>
<td>117.85</td>
<td>117.85</td>
<td>117.85</td>
<td>117.85</td>
</tr>
<tr>
<td>rc207.4</td>
<td>6</td>
<td>119.64</td>
<td>119.64</td>
<td>119.64</td>
<td>119.64</td>
</tr>
<tr>
<td>rc202.2</td>
<td>14</td>
<td>304.14</td>
<td>304.14</td>
<td>304.14</td>
<td>304.14</td>
</tr>
<tr>
<td>rc205.1</td>
<td>14</td>
<td>343.21</td>
<td>343.21</td>
<td>343.21</td>
<td>343.21</td>
</tr>
<tr>
<td>rc203.4</td>
<td>15</td>
<td>314.29</td>
<td>314.29</td>
<td>314.29</td>
<td>314.29</td>
</tr>
<tr>
<td>rc203.1</td>
<td>19</td>
<td>453.48</td>
<td>453.48</td>
<td>453.48</td>
<td>453.48</td>
</tr>
<tr>
<td>rc201.1</td>
<td>20</td>
<td>444.54</td>
<td>444.54</td>
<td>444.54</td>
<td>444.54</td>
</tr>
<tr>
<td>rc204.3</td>
<td>24</td>
<td>455.03</td>
<td>455.03</td>
<td>455.03</td>
<td>455.03</td>
</tr>
<tr>
<td>rc206.3</td>
<td>25</td>
<td>574.42</td>
<td>574.42</td>
<td>574.42</td>
<td>574.42</td>
</tr>
<tr>
<td>rc201.2</td>
<td>26</td>
<td>711.54</td>
<td>711.54</td>
<td>711.54</td>
<td>711.54</td>
</tr>
<tr>
<td>rc201.4</td>
<td>26</td>
<td>793.64</td>
<td>793.64</td>
<td>793.64</td>
<td>793.64</td>
</tr>
<tr>
<td>rc205.2</td>
<td>27</td>
<td>755.93</td>
<td>755.93</td>
<td>755.93</td>
<td>755.93</td>
</tr>
<tr>
<td>rc202.4</td>
<td>28</td>
<td>793.03</td>
<td>793.03</td>
<td>800.18</td>
<td>793.03</td>
</tr>
<tr>
<td>rc205.4</td>
<td>28</td>
<td>760.47</td>
<td>760.47</td>
<td>765.38</td>
<td>760.47</td>
</tr>
</tbody>
</table>
## Experiment results (3)

<table>
<thead>
<tr>
<th>Problem</th>
<th># cities</th>
<th>State of the art</th>
<th>NMC_EK score</th>
<th>NRPA score</th>
<th>NRPA_EK score</th>
</tr>
</thead>
<tbody>
<tr>
<td>rc202.3</td>
<td>29</td>
<td>837.72</td>
<td><strong>837.72</strong></td>
<td>839.58</td>
<td>839.58</td>
</tr>
<tr>
<td>rc208.2</td>
<td>29</td>
<td>533.78</td>
<td>536.04</td>
<td>537.74</td>
<td><strong>533.78</strong></td>
</tr>
<tr>
<td>rc207.2</td>
<td>31</td>
<td>701.25</td>
<td>707.74</td>
<td>702.17</td>
<td><strong>701.25</strong></td>
</tr>
<tr>
<td>rc201.3</td>
<td>32</td>
<td>790.61</td>
<td><strong>790.61</strong></td>
<td>796.98</td>
<td><strong>790.61</strong></td>
</tr>
<tr>
<td>rc204.2</td>
<td>33</td>
<td>662.16</td>
<td>675.33</td>
<td>673.89</td>
<td>664.38</td>
</tr>
<tr>
<td>rc202.1</td>
<td>33</td>
<td>771.78</td>
<td>776.47</td>
<td>775.59</td>
<td>772.18</td>
</tr>
<tr>
<td>rc203.2</td>
<td>33</td>
<td>784.16</td>
<td><strong>784.16</strong></td>
<td>784.16</td>
<td><strong>784.16</strong></td>
</tr>
<tr>
<td>rc207.3</td>
<td>33</td>
<td>682.40</td>
<td>687.58</td>
<td>688.50</td>
<td><strong>682.40</strong></td>
</tr>
<tr>
<td>rc207.1</td>
<td>34</td>
<td>732.68</td>
<td>743.29</td>
<td>743.72</td>
<td><strong>738.74</strong></td>
</tr>
<tr>
<td>rc205.3</td>
<td>35</td>
<td>825.06</td>
<td>828.27</td>
<td>828.36</td>
<td><strong>825.06</strong></td>
</tr>
<tr>
<td>rc208.3</td>
<td>36</td>
<td>634.44</td>
<td>641.17</td>
<td>656.40</td>
<td>650.49</td>
</tr>
<tr>
<td>rc203.3</td>
<td>37</td>
<td>817.53</td>
<td>837.72</td>
<td>820.93</td>
<td><strong>817.53</strong></td>
</tr>
<tr>
<td>rc206.2</td>
<td>37</td>
<td>828.06</td>
<td>839.18</td>
<td>829.07</td>
<td><strong>828.06</strong></td>
</tr>
<tr>
<td>rc206.4</td>
<td>38</td>
<td>831.67</td>
<td>859.07</td>
<td>831.72</td>
<td><strong>831.67</strong></td>
</tr>
<tr>
<td>rc208.1</td>
<td>38</td>
<td>789.25</td>
<td>797.89</td>
<td>799.24</td>
<td><strong>793.60</strong></td>
</tr>
<tr>
<td>rc204.1</td>
<td>46</td>
<td>868.76</td>
<td>899.79</td>
<td>883.85</td>
<td><strong>880.89</strong></td>
</tr>
</tbody>
</table>
Outline

- Traveling Salesman Problem with Time Windows
- Nested Monte-Carlo Algorithm
- Nested Rollout Policy Adaptation
- Experiments
- Conclusion
Conclusion

- **Results**
  - Efficient algorithm (77% of SOTA scores for NRPA_EK)
  - Promising results with no/few domain knowledge.
  - Expert knowledge is always helpful
  - Difficulties when the number of nodes becomes too large.

- **Current work**
  - Beam NRPA
  - Local optima issues?
Thank you