Cooperative Logistics: Can Artificial Intelligence Enable Trustworthy Cooperation at Scale?

Stephen Mak[1] **Tim Pearce**[2] **Matthew Macfarlane**[3,4] **Liming Xu**[1]
Michael Ostroumov[5] **Alexandra Brintrup**[1,6]
[1] University of Cambridge [2] Microsoft Research [3] University of Amsterdam
[4]InstaDeep [5] Value Chain Lab [6] The Alan Turing Institute

Abstract

Cooperative Logistics studies the setting where logistics companies pool their resources together to improve their individual performance. Prior literature suggests carbon savings of approximately 22%. If attained globally, this equates to 480,000,000 tonnes of CO_2. Whilst well-studied in operations research – industrial adoption remains limited due to a lack of trustworthy cooperation. A key remaining challenge is *fair* and *scalable* gain sharing (i.e., how much should each company be fairly paid?). This paper introduces the novel algorithmic challenges that Cooperative Logistics offers AI, and novel applications of AI towards Cooperative Logistics. We further present findings from our initial experiments.

1 Introduction

The transportation industry[1] emitted 7.1 billion tonnes of CO_2 ($GtCO_2$) globally in 2020 [1, p. 64], or 13% of global greenhouse gas emissions [2]. To limit the global average temperature increase to 1.5°C, the Intergovernmental Panel of Climate Change (IPCC) states that the transportation industry must decrease their CO_2 emissions to 2.6 $GtCO_2$ by 2050 [3].

Road freight transportation emitted 2.2 $GtCO_2$ globally in 2020[2]. Yet, under currently implemented and announced policies, the International Transport Forum expects road freight transportation emissions to *increase* to 2.5 $GtCO_2$ in 2050 – single-handedly exhausting almost all of the transportation industry's carbon budget. In addition, road freight transportation is seen as one of the most difficult industries to decarbonise [5]. This is due to the sheer weights and distances travelled. In 2020, global road freight activity was at 26.8 trillion tonne-kilometers[3]. This is expected to double by 2050 due to an increasing global population and global economic growth [1].

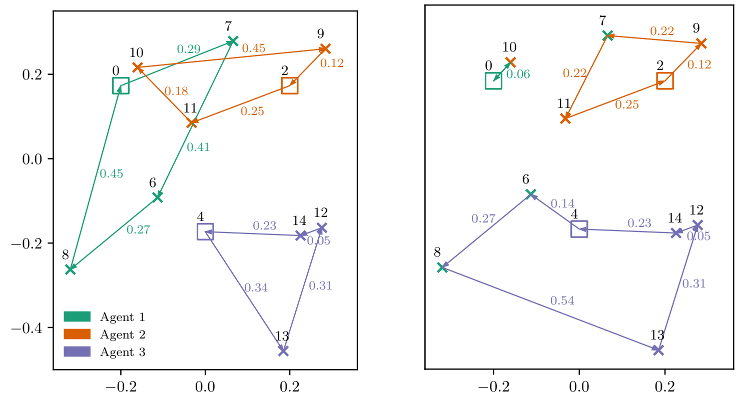
One approach to decarbonise road freight transportation is through cooperation. Cooperative vehicle routing studies the setting where delivery companies share their delivery information and perform deliveries on behalf of one another. An illustration can be found in Figure 1. Cooperative vehicle routing has been studied in operations research for at least two decades, with estimated cost savings between 4-46% [6, 7, 8, 9, 10]. Cooperative vehicle routing results in reduced distance travelled by trucks, thus reducing carbon emissions and road congestion. In addition, delivery companies can increase their revenue, improve customer service, and gain market share [9]. Surprisingly, despite the abundance of perceived benefits acknowledged by industry, government, and academia – cooperative vehicle routing has still not seen widespread industrial traction [8].

[1]Transportation involves the movement of people and goods across air, road, rail and water.

[2]Due to COVID-19, global freight activity dropped by 4%, but global freight emissions only dropped by 1% due to increased urban deliveries [4, p. 194].

[3]A tonne-kilometer is when a tonne of goods is transported one kilometer.

Computational Sustainability Workshop at NeurIPS 2023.



(a) Before cooperation
Total Cost = 3.35

(b) After cooperation
Total Cost = 2.47

Figure 1: An example of three companies (denoted by colours) before and after cooperation. Through cooperation, the companies are able to attain more efficient routes, reducing total cost by 0.88 or 26%. Squares denote depots. Crosses denote customer locations. Node indices (arbitrary) are denoted in black, with route costs given along edges.

A key remaining barrier is the lack of *scalable* and *trustworthy* algorithms for gain sharing, i.e., if two or more companies cooperate, how much should everyone be paid? Fairness is important to facilitate trustworthy cooperation. Cooperative game theory allows us to study fair gain sharing in a principled manner and has been widely applied to cooperative logistics [6, 11]. The Shapley value is widely considered as a fair solution concept; however, computing this value for cooperative vehicle routing involves solving 2^n vehicle routing problems, each of which is NP-hard. Consequently, the prior literature on fair cooperative vehicle routing usually studies settings with at most 6 agents [10]. However, scalability is particularly important as the logistics industry is highly fragmented. In the UK, there are 60,000 registered carriers in 2022 [12]. In the EU, there are 1 million registered carriers in 2020 [13]. Moreover, prior literature establishes that the network externalities[4] are significant and a carbon reduction of 22.1% could be achieved in a setting with 1000 cooperating carriers [14]. To achieve this scalability, they pose the problem as a coalition structure generation problem and propose a correlation-based heuristic to identify fruitful coalitions. Our contribution is that we pose coalition structure generation as a deep reinforcement learning problem and our initial experiments show promise at small scales. We further introduce a neural reward model to generate large volumes of training experience. However, more importantly, there is still a vast array of interesting algorithmic challenges to focus on within cooperative logistics and we outline avenues for future research which we hope would be of interest to the Trustworthy Cooperative AI community [15, 16].

2 Problem Formulation and Experimental Setup

We wish to solve the problem with a set $N = \{1, \dots, n\}$ companies cooperating, where in our setting we hope to ultimately scale to at least $n = 1000$ companies. A *coalition* C is a subset of the agents (or companies) N, i.e., $C \subseteq N$. A *coalition structure* CS over N is a collection of coalitions $CS = \{C^1, \dots, C^k\}$ such that $\cup_{j=1}^{k} C^j = N$ and $C^i \cap C^j = \varnothing$ for any $i, j \in \{1, \dots, k\} : i \neq j$ [17]. The *characteristic function* $v : 2^N \to \mathbb{R}_{\geq 0}$ maps each coalition $C \subseteq N$ to a real number $v(C)$ which is called the *value* of the coalition C. In Figure 1, $v(\{1, 2, 3\}) = 3.35 - 2.47 = 0.88$. Note, to compute the value of a single coalition involves solving a multi-depot vehicle routing problem (NP-hard). The *social welfare* is the sum of the value of each coalition in the coalition structure, i.e. $\sum_{C \in CS} v(C)$. Coalition structure generation aims to find the coalition structure CS that maximises social welfare. Note that the number of possible coalition structures is the Bell number [18], B_n, where $B_{1000} = 2.98 \times 10^{1927}$. To manage this complexity, as well as to enable tractable gain sharing,

[4]Network externalities, or the network effect, is where the value of a good or service increases with an increasing number of users, such as social media platforms.

we enforce that each coalition has size $|C| = M$, in our case, $M = 3$. After a coalition structure is obtained, each of the (much smaller) sub-coalitions would perform individual fair gain sharing. To calculate the Shapley value without coalition structure generation would require solving $2^{N=1000}$ NP-hard VRPs. Using coalition structure generation, this can be seen as a hierarchical approach and now "only" requires $\lfloor \frac{N}{M} \rfloor \cdot 2^M$ NP-hard VRPs to be solved where $M << N$. Alternatively, each sub-coalition could have multi-agent reinforcement learning agents bargain with each other to agree on a fair gain share, resulting in significantly reduced run-times [19, 20, 21].

Data and Experimental Setup: We follow the same experimental setup as in [14] and compare our RL approach to theirs. We collaborate with industry to gather real-world data from the Food & Beverages industry in the UK. The dataset describes 34,692 shipments with 6 unique warehouses (origins) and 976 unique customer locations (destinations) which occupy a varying number of pallets q in a truck. Each truck is assumed to have a capacity Q of 12 pallets. Each shipment also has a price p which the carrier receives for performing that delivery. A training instance is procedurally generated by sub-sampling from the dataset. Each truck is randomly assigned m shipments, in our case $m = 4$. Note that due to the capacity Q, each truck may not necessarily be able to satisfy all four shipments and some shipments may need to be dropped. Thus, the state space has $m \times n$ rows (or shipments), with 10 features to describe each shipment. Due to the large state space, flattening the state and using multi-layer perceptrons is not scalable. In addition, there is permutation invariance in the state, which we can leverage through the use of transformers [22].

At each timestep t, the RL agent selects M companies to form a coalition. We apply invalid action masking to mask out previously selected companies. The value of those M selected companies, $v(C^t)$, is estimated via a neural reward model. An episode continues until all companies are assigned to a coalition and a feasible coalition structure is obtained. The total (discounted) return the RL agent receives is thus the social welfare, which it aims to maximise. We use Proximal Policy Optimisation (PPO) [23] which we implement in TensorFlow [24]. We compare our approach against the previous state of the art [14] and also compare against an agent that selects coalitions uniformly at random.

Neural Reward Model: To calculate the value of a coalition $v(C)$ requires solving a multi-depot vehicle routing problem either exactly or approximately, with solvers such as Gurobi or Google OR Tools. However, to train our RL agent would require a large volume of training examples, especially when using on-policy policy gradient methods. Furthermore, due to the high variance in RL, larger batch sizes are preferable. Using traditional solvers would not be able to generate sufficient volumes of training data. Instead, we train a neural network to predict the value of a coalition $v(C)$. We sub-sample from our real-world dataset and select coalitions C uniformly at random. We then train a transformer to predict the value of a coalition $v(C)$. We find that multi-task learning also helps in this scenario: to calculate $v(C)$ also requires knowing the revenues, costs, profits and distances travelled for each agent before and after cooperation. Therefore, we predict these auxiliary targets as well and minimise mean squared error on all heads.

3 Results and Discussion

The primary metric we train the agent to optimise is the (monetary) social welfare. In Table 1, we present the results of our RL agent trained on instances with 6, 9, 12 and 15 companies as this is early work in progress. Future work will of course investigate additional opportunities to improve performance and scale to the real-world size of at least 1,000 companies. For 6 and 9 agents, our RL method is able to outperform the previous state of the art [14] with statistical significance. For 12 agents, our RL method performs on par with the prior method; however, for 15 agents, our RL method performs worse than previous state of the art. We further present results for the carbon reduction in Appendix A due to space limitations. Surprisingly, even in the 6 and 9 agent setting, in Table 2 we accept the null hypothesis that our RL agent saves as much carbon emissions as prior state of the art, even though our approach outperforms in terms of social welfare. It is important that future work is cognizant that solely optimising for social welfare may not directly translate to carbon emissions. However, this should not be interpreted as a disadvantage of our proposed method. Cooperation is highly desirable to reduce carbon emissions, yet has seen limited success to date. With increased social welfare, this presents increased opportunities to further incentivise cooperation. Future work should investigate the most appropriate method to balance social welfare with sustainable development goals.

Table 1: Comparison of methods on 1,000 paired test instances

Method	Social welfare for n agents $n=6$	$n=9$	$n=12$	$n=15$
Random Agent	1057.2	1590.3	2123.5	2641.6
Anonymous et al. [14]	1101.9	1659.6	2234.1	**2800.1****
RL Agent (Ours)	**1142.0*****	**1705.0*****	**2236.9**	2759.0

H_0: difference in performance between RL agent and Anonymous et al. [14] is 0.
*** = significant at 0.001 level. ** = significant at 0.01 level.

There is also ample opportunity to improve our method. Firstly, we believe that the nature of this problem is highly amenable to curriculum learning. We hypothesise that, due to our deliberate neural network architecture, a transformer-based agent trained on e.g. $n = 6$ agents should generalise well, and learn useful representations that positively transfer out of distribution to settings with more (or less) agents. We see slight evidence for this in similar work to ours [25, 26]. Secondly, whilst we train our agents tabula rasa, it may be desirable to pre-train the agents on similar tasks. For example, the learnt representations from our neural reward model may have learnt useful features that could be transferable to the RL setting. Whilst not as elegant, this is commonplace in supervised learning, and used in most of the large-scale RL successes [27, 28, 29], drawing recent interest [30]. Furthermore, search-based methods could help significantly guide exploration. Finally, this problem setting is interesting due the order invariance of the actions: it does not matter in which order the coalitions are selected in a coalition structure. How can we leverage this invariance?

4 Conclusion and Future Work

Our initial findings suggest that the use of deep reinforcement learning, transformers, and a neural reward model for coalition structure generation to be effective in small scale scenarios. We believe this is due to the transformer's ability to attend over the entire state space without losing information, which was required in prior literature. A natural future research direction is to improve our approach to scale to real-world settings. However, outside of this, we believe there is still a vast array of interesting algorithmic challenges remaining within cooperative logistics and cooperative supply chains in general and believe that cross-fertilisation of AI and operations research to be highly fruitful.

What AI can offer Cooperative Logistics: Cooperative game theory is desirable for its principled approach towards fairly incentivising rational agents. However, this is computationally challenging due to the central assumption of having access to the characteristic function, which would require an exponential number of NP-hard vehicle routing problems to be solved. Instead, the use of high capacity neural networks allows us to remove this assumption for both coalition structure generation and fair gain allocation. Moreover, the use of deep MARL for negotiation allows us to obtain fair incentivisation with significantly reduced run-time [19, 14, 21, 31, 32]. In addition, can Explainable AI help explain to human transportation managers why proposed profit allocations are fair?

What Cooperative Logistics can offer AI: Cooperative game theory is currently not sufficient for cooperative logistics. Cooperative game theory assumes that cooperation only occurs once. In the real world, cooperation involves repeated interactions throughout time. If a company tries to take (unfair) advantage over its competitor, this may affect its *reputation*. Moreover, not all companies are equal – some have more market power than others. How do we account for these power dynamics? We believe that evolutionary game theory and deep MARL could be useful to study this setting. The use of natural language in autonomous negotiation would also be exciting to understand agent reasoning.

Time is ticking to tackle climate change. Road freight transportation is certainly not on track to meet international climate objectives [1]. Industry, government and academia already acknowledge the benefits of cooperative logistics [10]. But trustworthy cooperation between companies is not easy, both computationally and in practice. Can AI enable trustworthy cooperation at scale?

Acknowledgments and Disclosure of Funding

This work was supported by the UK Engineering and Physical Sciences Research Council (EPSRC) grant on "Intelligent Systems for Supply Chain Automation" under Grant Number 2275316.

This work was performed using resources provided by the Cambridge Service for Data Driven Discovery (CSD3) operated by the University of Cambridge Research Computing Service (www.csd3.cam.ac.uk), provided by Dell EMC and Intel using Tier-2 funding from the Engineering and Physical Sciences Research Council (capital grant EP/T022159/1), and DiRAC funding from the Science and Technology Facilities Council, United Kingdom (www.dirac.ac.uk).

We thank the Supply Chain Artificial Intelligence Lab (SCAIL) for their insightful discussions regarding early drafts of this paper.

References

[1] ITF. *ITF Transport Outlook 2023*. ITF Transport Outlook. OECD, May 2023.

[2] M Crippa, D Guizzardi, F Pagani, M Banja, M Muntean, E Schaaf, W Becker, F Monforti-Ferrario, R Quadrelli, A Risquez Martin, P Taghavi-Moharamli, J Köykkä, G Grassi, S Rossi, J Brandao De Melo, D Oom, A Branco, J San-Miguel, and E Vignati. GHG emissions of all world countries. Publications Office of the European Union doi:10.2760/953332, European Commission, Luxembourg, 2023.

[3] IPCC. *Global Warming of 1.5 C: IPCC Special Report on Impacts of Global Warming of 1.5 C above Pre-industrial Levels in Context of Strengthening Response to Climate Change, Sustainable Development, and Efforts to Eradicate Poverty*. Cambridge University Press, 1 edition, June 2022.

[4] ITF. *ITF Transport Outlook 2021*. Organisation for Economic Co-operation and Development, Paris, 2021.

[5] Jeffrey Sachs, Laurence Tubiana, Emmanuel Guerin, Henri Waisman, Carl Mas, Michel Colombier, and Guido Schmidt-Traub. Pathways to deep decarbonization: 2014 report. Technical report, Sustainable Development Solutions Network and Institute for Sustainable Development and International Relations, 2014.

[6] Mario Guajardo and Mikael Rönnqvist. A review on cost allocation methods in collaborative transportation. *International Transactions in Operational Research*, 23(3):371–392, 2016. Publisher: Wiley.

[7] Margaretha Gansterer and Richard F. Hartl. Collaborative vehicle routing: A survey. *European Journal of Operational Research*, 268(1):1–12, 2018. Publisher: Elsevier BV.

[8] William Ferrell, Kimberly Ellis, Phil Kaminsky, and Chase Rainwater. Horizontal collaboration: opportunities for improved logistics planning. *International Journal of Production Research*, 58(14):4267–4284, 2020. Publisher: Informa UK Limited.

[9] Frans Cruijssen, Martine Cools, and Wout Dullaert. Horizontal cooperation in logistics: Opportunities and impediments. *Transportation Research Part E: Logistics and Transportation Review*, 43(2):129–142, 2007. Publisher: Elsevier BV.

[10] Frans Cruijssen. *Cross-Chain Collaboration in Logistics: Looking Back and Ahead*. International Series in Operations Research and Management Science. Springer, January 2020.

[11] Frederik Schulte, Eduardo Lalla-Ruiz, Silvia Schwarze, Rosa Gonzalez Ramirez, and Stefan Voss. *Scalable Core and Shapley Value Allocation Methods for Collaborative Transportation*. July 2019.

[12] Office for National Statistics. UK business: activity, size and location - Office for National Statistics, September 2022.

[13] Eurostat. Annual detailed enterprise statistics for services (NACE Rev. 2 H-N and S95): SBS_na_1a_se_r2, 2020.

[14] A Anonymous. *Journal Paper Under Review*, 2023.

[15] Allan Dafoe, Edward Hughes, Yoram Bachrach, Tantum Collins, Kevin R. McKee, Joel Z. Leibo, Kate Larson, and Thore Graepel. Open Problems in Cooperative AI. *arXiv:2012.08630 [cs]*, December 2020. arXiv: 2012.08630.

[16] Alexandra Brintrup, George Baryannis, Ashutosh Tiwari, Svetan Ratchev, Giovanna Martinez-Arellano, and Jatinder Singh. *Trustworthy, responsible, ethical AI in manufacturing and supply chains: synthesis and emerging research questions*. May 2023.

[17] Georgios Chalkiadakis, Edith Elkind, and Michael Wooldridge. *Computational Aspects of Cooperative Game Theory (Synthesis Lectures on Artificial Inetlligence and Machine Learning)*. Morgan & Claypool Publishers, 1st edition, 2011.

[18] E. T. Bell. Exponential Numbers. *The American Mathematical Monthly*, 41(7):411–419, August 1934. Publisher: Taylor & Francis _eprint: https://doi.org/10.1080/00029890.1934.11987615.

[19] Stephen Mak, Liming Xu, Tim Pearce, Michael Ostroumov, and Alexandra Brintrup. Coalitional Bargaining via Reinforcement Learning: An Application to Collaborative Vehicle Routing. In *NeurIPS Cooperative AI Workshop*, October 2021.

[20] A. Anonymous. *Journal Paper Under Review*, 2023.

[21] Yoram Bachrach, Richard Everett, Edward Hughes, Angeliki Lazaridou, Joel Z. Leibo, Marc Lanctot, Michael Johanson, Wojciech M. Czarnecki, and Thore Graepel. Negotiating team formation using deep reinforcement learning. *Artificial Intelligence*, 288:103356, November 2020.

[22] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention Is All You Need. *arXiv:1706.03762 [cs]*, December 2017. arXiv: 1706.03762.

[23] John Schulman, Filip Wolski, Prafulla Dhariwal, Alec Radford, and Oleg Klimov. Proximal Policy Optimization Algorithms. *arXiv:1707.06347 [cs]*, August 2017. arXiv: 1707.06347.

[24] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dandelion Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: Large-Scale Machine Learning on Heterogeneous Systems, 2015.

[25] Daphne Cornelisse, Thomas Rood, Mateusz Malinowski, Yoram Bachrach, and Tal Kachman. Neural Payoff Machines: Predicting Fair and Stable Payoff Allocations Among Team Members. arXiv, August 2022. arXiv:2208.08798 [cs, econ].

[26] Wouter Kool, Herke van Hoof, and Max Welling. Attention, Learn to Solve Routing Problems! *arXiv:1803.08475 [cs, stat]*, February 2019. arXiv: 1803.08475.

[27] David Silver, Aja Huang, Chris J. Maddison, Arthur Guez, Laurent Sifre, George van den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, Sander Dieleman, Dominik Grewe, John Nham, Nal Kalchbrenner, Ilya Sutskever, Timothy Lillicrap, Madeleine Leach, Koray Kavukcuoglu, Thore Graepel, and Demis Hassabis. Mastering the game of Go with deep neural networks and tree search. *Nature*, 529(7587):484–489, January 2016.

[28] Oriol Vinyals, Igor Babuschkin, Wojciech M. Czarnecki, Michaël Mathieu, Andrew Dudzik, Junyoung Chung, David H. Choi, Richard Powell, Timo Ewalds, Petko Georgiev, Junhyuk Oh, Dan Horgan, Manuel Kroiss, Ivo Danihelka, Aja Huang, Laurent Sifre, Trevor Cai, John P. Agapiou, Max Jaderberg, Alexander S. Vezhnevets, Rémi Leblond, Tobias Pohlen, Valentin Dalibard, David Budden, Yury Sulsky, James Molloy, Tom L. Paine, Caglar Gulcehre, Ziyu Wang, Tobias Pfaff, Yuhuai Wu, Roman Ring, Dani Yogatama, Dario Wünsch, Katrina Mckinney, Oliver Smith, Tom Schaul, Timothy Lillicrap, Koray Kavukcuoglu, Demis Hassabis, Chris Apps, and David Silver. Grandmaster level in StarCraft II using multi-agent reinforcement learning. *Nature*, 575(7782):350–354, 2019. Publisher: Springer Science and Business Media LLC.

[29] OpenAI, Christopher Berner, Greg Brockman, Brooke Chan, Vicki Cheung, Przemysław Dębiak, Christy Dennison, David Farhi, Quirin Fischer, Shariq Hashme, Chris Hesse, Rafal Józefowicz, Scott Gray, Catherine Olsson, Jakub Pachocki, Michael Petrov, Henrique P. d O. Pinto, Jonathan Raiman, Tim Salimans, Jeremy Schlatter, Jonas Schneider, Szymon Sidor, Ilya Sutskever, Jie Tang, Filip Wolski, and Susan Zhang. Dota 2 with Large Scale Deep Reinforcement Learning. *arXiv:1912.06680 [cs, stat]*, December 2019. arXiv: 1912.06680.

[30] Rishabh Agarwal, Max Schwarzer, Pablo Samuel Castro, Aaron Courville, and Marc G. Bellemare. Reincarnating Reinforcement Learning: Reusing Prior Computation to Accelerate Progress, October 2022. arXiv:2206.01626 [cs, stat].

[31] Zun Li, Marc Lanctot, Kevin R. McKee, Luke Marris, Ian Gemp, Daniel Hennes, Paul Muller, Kate Larson, Yoram Bachrach, and Michael P. Wellman. Combining Tree-Search, Generative Models, and Nash Bargaining Concepts in Game-Theoretic Reinforcement Learning, February 2023. arXiv:2302.00797 [cs].

[32] János Kramár, Tom Eccles, Ian Gemp, Andrea Tacchetti, Kevin R. Mckee, Mateusz Malinowski, Thore Graepel, and Yoram Bachrach. Negotiation and honesty in artificial intelligence methods for the board game of Diplomacy. *Nature Communications*, 13(1), 2022. Publisher: Springer Science and Business Media LLC.

[33] UK DESNZ and UK BEIS. Greenhouse gas reporting: conversion factors 2021, January 2022.

A Carbon Emissions Calculations and Results

The UK's Department for Energy Security and Net Zero, and Department for Business, Energy & Industrial Strategy estimates that the average laden (loaded) heavy goods vehicle, averaged across all types of heavy goods vehicles emits 107.5 gCO_2-eq per tonne kilometer [33]. Since we know the weight of each shipment, combined with the distances travelled, we can estimate the CO_2 emissions before cooperation and after cooperation, resulting in the below table. Of interest, note that the CO_2 reduction in general increases with the number of cooperating companies, thus demonstrating the network externalities present in this system. Thus, it is desirable to obtain an algorithm that can scale to a large number of companies. In the work of [14], they show that in a system with 1,002 companies, a carbon reduction of 22.1% could be achieved.

Table 2: Comparison of methods on 1,000 test instances

	$n = 6$		$n = 9$		$n = 12$		$n = 15$	
Method	CO_2 Reduction (%)	95% CI	CO_2 Reduction (%)	95% CI	CO_2 Reduction (%)	95% CI	CO_2 Reduction (%)	95% CI
Random Agent	9.12	[7.52, 10.7]	8.64	[7.36, 9.92]	8.77	[7.70, 9.85]	8.48	[7.48, 9.49]
Anonymous et al. [14]	10.8	[9.23, 12.4]	**12.9**	[11.6, 14.2]	**14.5*****	[13.4, 15.6]	**14.6*****	[13.7, 15.6]
RL Agent (Ours)	**11.1**	[9.45, 12.7]	12.2	[10.9, 13.5]	11.7	[10.6, 12.8]	11.2	[10.2, 12.1]

H_0: difference in performance between RL agent and Anonymous et al. [14] is 0.
*** = significant at 0.001 level. ** = significant at 0.01 level.