Available online at www.sciencedirect.com

ScienceDirect

Cognitive Systems Research 52 (2018) 387–399

Cognitive Systems RESEARCH

www.elsevier.com/locate/cogsys

Human behaviour in the Euclidean Travelling Salesperson Problem: Computational modelling of heuristics and figural effects

Action editor: Ronaldo Vigo

Markos Kyritsis [a,*], Stephen R. Gulliver [b], Eva Feredoes [c], Shahab Ud Din [a]

[a] *Computer and Information Science Department, Higher Colleges of Technology, Sharjah, United Arab Emirates*
[b] *Henley Business School, Business Informatics Systems and Accounting, University of Reading, Whiteknights Road, RG6 6UD, United Kingdom*
[c] *School of Psychology and Clinical Language Sciences, University of Reading, Whiteknights Road, RG6 6AL, United Kingdom*

Received 16 February 2018; received in revised form 13 June 2018; accepted 20 July 2018
Available online 29 July 2018

Abstract

The Travelling Salesperson Problem (TSP) describes a situation where an imaginary individual wishes to visit multiple cities once before returning to his/her own city. This type of problem is known as a nondeterministic polynomial (NP) hard problem, since the factorial number of solutions results in it being impractical to solve using exhaustive processing. Interestingly, when presented as a Euclidean graph (i.e., ETSP), humans identify near optimal solutions almost effortlessly, despite billions of possible tours. In this study, we consider human processing of the ETSP, and introduce the reader to a number of factors that literature proposes as impacting human performance. We hypothesise that: (i) human ETSP activity may be modelled by considering the quotient relationship between node-to-node and node-to-centroid distances; and (ii) consideration of figural effects can optimise automated TSP solution generation. In this paper human processing based heuristics are developed, i.e. replacing the cost function within the nearest neighbour algorithm, to guide node selection. Results showed that the quotient relationship between node-to-node and node-to-centroid distances can be used to closely model average human performance, across a range of ETSP graphs. Interestingly, however, additional consideration of graph figural effects (e.g. distance between boundary points in the convex hull, standard deviation of distances between nodes that make up the convex hull, and number of nodes in the convex hull) results in significantly improved tour costs.

Keywords: Travelling Salesperson Problem; Computational model; Heuristics

1. Introduction

A salesperson wants to visit every city on a map once and once only, before returning home – the classic Travelling Salesperson Problem (TSP). Even though this might sound like a simple problem, if we assume there is a fixed starting and finishing point, there are $(n - 1)!/2$ possible solutions – where n is the number of nodes, i.e., cities. Therefore, calculating the optimal route between a large number of nodes can quickly become computationally expensive (Dry, Lee, Vickers, & Hughes, 2006). Case in point, in 1962 Procter and Gamble offered US$10,000 to the person who could identify the shortest route between 33 U.S. cities (1.32×10^{35} possible routes), a problem that they – at the time – estimated would take 417 billion trillion years to exhaustively computationally analyse (Applegate, Bixby, Chvatal, & Cook, 2007). The TSP is part of a wider

* Corresponding author.
E-mail address: mkyritsis@hct.ac.ae (M. Kyritsis).

https://doi.org/10.1016/j.cogsys.2018.07.027
1389-0417/

set of problems, known as nondeterministic polynomial hard (NP-hard) problems. NP-hard problems occur in a wide range of industrial domains, including: genome sequencing (Agarwala, Applegate, Maglott, Schuler, & Schaffer, 2000), analysis of crystal structures in X-ray diffraction (Bland & Shallcross, 1987), logistics (Dallari, Marchet, & Ruggeri, 2000), the selection of routes for e.g., school buses and delivery of meals to the elderly (Schrijver, 2003), computer wiring, etc. (for an overview of applications, see Matai, Singh, & Mittal, 2010; MacGregor & Chu, 2011).

Since use of brute force is practically impossible, practitioners managing NP-hard problems often apply near-optimal heuristic methods, which produce close to optimal solutions at a fraction of the processing cost. Despite the computational complexity of the NP-hard problems, when a TSP graph is visually represented as a Euclidean graph (i.e., a ETSP), the average human can produce near optimal solutions in near-linear time (Dry et al., 2006; Graham, Joshi, & Pizlo, 2000); producing tour costs that are approximately 1% more than optimal for problems where the number of nodes is between 10 and 20 (MacGregor & Ormerod, 1996), and approximately 11% more than optimal for problems where the number of nodes is 120 (Dry et al., 2006). How do humans process ESTP graphs so efficiently, and can we learn from this?

MacGregor and Ormerod (1996) suggest that when solving the ETSP, humans apply a global-to-local heuristic; i.e. starting at the whole and subsequently focusing on each node. The two approaches discussed in this section, are the convex hull hypothesis and hierarchical course-to-fine heurists. The convex hull hypothesis involves initially forming an imaginary perimeter i.e. a convex hull, around the boundary nodes (see Fig. 1). Once the convex hull is in place, individuals then sequentially insert local internal

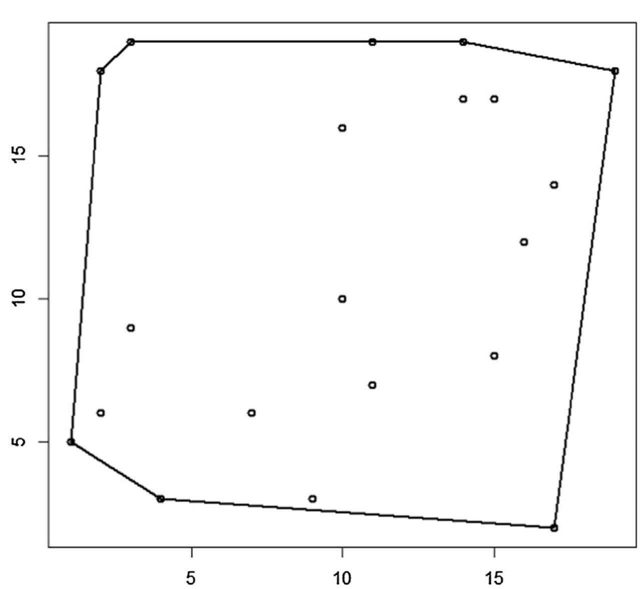

Fig. 1. Convex hull of a graph.

nodes in either a clockwise or a counter clockwise direction (MacGregor, Chronicle, & Ormerod, 2004). Empirical ETSP literature shows that various convex hull properties, e.g. the number of boundary nodes, angles of lines in the hull, and the spread of nodes from the centroid, correlate significantly with individual ETSP performance (Kyritsis, Gulliver, & Feredoes, 2017; Vickers, Lee, Dry, & Hughes, 2003). Hierarchical course-to-fine heuristics (e.g., Graham et al., 2000) apply a global-to-local approach to model human performance. For example, the Foveating Pyramidal Model (FPM), developed by Pizlo et al. (2006), uses graph bisection in combination with Gaussian filtering to 'emulate' the physiological processes involved in moving items from peripheral to foveal vision. From the convex hull hypothesis we hypothesise, within this paper, that human ETSP activity may by modelled partly by considering the quotient relationship between node-to-node and node-to-centroid distances.

In contrast to the global-to-local theorists, other researchers apply a local-to-global approach, i.e. identifying the candidate node based on the low level ETSP graph properties. In support of the local-to-global theorists, research shows that geometric properties, not necessarily related to the convex hull, also impact performance ETSP performance, e.g. node randomness, node clustering, and node regularity (Dry, Preiss, & Wagemans, 2012). Interestingly both global-to-local and local-to-global theorists all agree that optimal ETSP solutions are absent of crossings (Graham et al., 2000; MacGregor & Ormerod, 1996). Humans actively seek to avoid crossing their own path; i.e. a local precedence visual search strategy named 'the crossing-avoidance hypothesis' (van Rooij, Stege, & Schactman, 2003). Theorists suggest that humans are 'trained' that tours with crossed lines are non-optimal. It is a matter of debate whether humans actively avoid crossing paths, or whether the act of creating a convex hull naturally results in graphs with a low number of crossings. Either way, van Rooij et al. (2003) found that crossings occur very infrequently in participant solutions, suggesting that minimising crossing avoidance is critical to human ETSP solving strategies, which is also necessary for optimal graph solutions.

Interestingly, Kyritsis, Gulliver, et al. (2017) showed that the layout of nodes, and the figural properties of the graph, can increase the occurrence of crossings; thus resulting in suboptimal paths. Figural effects, in fact, have been shown to impact the viewer's ability to perceptually process Gestaltian effects (e.g., good continuity), with variation existing in the impact that such effects have between individuals (MacGregor et al., 2004). Vickers et al. (2003) also reported that human performance degrades as the number of nodes making up the convex hull increases. Moreover, Vickers et al. (2003) reported that the number of potential crossings, a measurement that was calculated by considering graph complexity and configuration, inversely impacted human performance. Despite supporting research, current heuristics do not include consideration of figural effects