

Computerizing Trail Making Test for long-term cognitive self-assessment

Trail Making
Test

83

Zhiwei Zeng

*Joint NTU-UBC Research Centre of Excellence in Active Living for the Elderly,
Nanyang Technological University, Singapore, Singapore
and Interdisciplinary Graduate School,
Nanyang Technological University, Singapore, Singapore*

Chunyan Miao

*Joint NTU-UBC Research Centre of Excellence in Active Living for the Elderly,
Nanyang Technological University, Singapore, Singapore
and School of Computer Science and Engineering,
Nanyang Technological University, Singapore, Singapore*

Cyril Leung

*Joint NTU-UBC Research Centre of Excellence in Active Living for the Elderly,
Nanyang Technological University, Singapore, Singapore, and*

Zhiqi Shen

*Joint NTU-UBC Research Centre of Excellence in Active Living for the Elderly,
Nanyang Technological University, Singapore, Singapore
and School of Computer Science and Engineering,
Nanyang Technological University, Singapore, Singapore*

Received 23 October 2016
Revised 11 January 2017
Accepted 15 January 2017

Abstract

Purpose – This paper aims to adapt and computerize the Trail Making Test (TMT) to support long-term self-assessment of cognitive abilities.

Design/methodology/approach – The authors propose a divide-and-combine (DAC) approach for generating different instances of TMT that can be used in repeated assessments with nearly no discernible practice effects. In the DAC approach, partial trails are generated separately in different layers and then combined to form a complete TMT trail.

Findings – The proposed approach was implemented in a computerized test application called *iTMT*. A pilot study was conducted to evaluate *iTMT*. The results show that the instances of TMT generated by the

© Zhiwei Zeng, Chunyan Miao, Cyril Leung and Zhiqi Shen. Published in the International Journal of Crowd Science. Published by Emerald Publishing. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licenses/by/4.0/legalcode>

This research is supported by the National Research Foundation, Prime Minister's Office, Singapore under its IDM Futures Funding Initiative; and the Interdisciplinary Graduate School Research Scholarship. The authors would like to thank Jun Ji for his contributions in application implementation and testing



DAC approach had an adequate level of difficulty. *iTMT* also achieved a stronger construct validity, higher test–retest reliability and significantly reduced practice effects than existing computerized tests.

Originality/value – The preliminary results suggest that *iTMT* is suitable for long-term monitoring of cognitive abilities. By supporting self-assessment, *iTMT* also can help to crowdsource the assessment processes, which need to be administered by healthcare professionals conventionally, to the patients themselves.

Keywords Telehealth, Self-assessment, Computerized cognitive assessment, Longitudinal assessment, Practice effect, Trail Making Test

Paper type Technical paper

1. Introduction

Both falls and dementia are major health concerns among the elderly. About one-third of the elderly aged over 65 years fall each year (Tinetti *et al.*, 1988). Falls also account for about 10 per cent of visits to hospital emergency departments among the elderly (Sattin, 1992). On the other hand, a new dementia case is reported every 3.2 seconds, and the cost of dementia is equivalent to about 1.1 per cent of global Gross Domestic Product (GDP) in 2015[1]. A decline in cognitive functions has been associated with increased risk of fall (Muir *et al.*, 2012) and identified as a precursor syndrome to dementia (Lyketsos *et al.*, 2002). Thus, monitoring changes in cognitive functions may be helpful for fall prevention, as well as early diagnosis and intervention of dementia.

Repeated assessments are required to track the changes in cognitive functions timely and effectively. The Trail Making Test (TMT) is one of the most frequently used neuropsychological tests for cognitive assessment (Butler *et al.*, 1991) due to its sensitivity, simplicity and ease of administration. The current version of TMT is adapted by Reitan (1955) (thereafter referred to as Reitan's TMT), which has only one instance and is administered using paper and pencil. However, the repeated use of Reitan's TMT in longitudinal assessment is severely limited by its high susceptibility to practice effects (Beglinger *et al.*, 2005). Practice effects refer to improvements in test performance that occur when a subject is retested on the same instance, or tested repeatedly on very similar ones. It is hard to segregate performance improvements due to practice effects from meaningful cognitive changes. This affects test accuracy and reliability. Increasing the test–retest interval may attenuate practice effects, but this may obscure the timely detection of meaningful cognitive changes (Buck *et al.*, 2008).

Some researchers have proposed to use alternative forms (Atkinson and Ryan, 2007) and mirror images (Wagner *et al.*, 2011) of TMT serially in consecutive test administrations to reduce practice effects. However, the number of equivalent alternative forms and mirror images of TMT is limited. Other researchers have proposed more systematic and divergent approaches for generating new instances of TMT (Vickers *et al.*, 1996; Vickers and Lee, 1998). Although their approaches could generate theoretically unlimited instances of TMT, the generated instances may be less difficult than Reitan's version – they may have shorter average trail length and less visual interference than the instance in Reitan's version. Consequently, the generated instances may have poorer diagnostic efficacy and less discriminating power to distinguish among different cognitive status. Moreover, as a cognitive assessment tool, the generated instances should be assessed for their validity and reliability; yet, to the best of our knowledge, none of these two approaches have been validated in user studies.

To generate TMTs that can be used in repeated assessments, we propose a divide-and-combine (DAC) approach – a systematic approach for generating instances of TMT which:

- (1) are sufficiently different from each other to reduce practice effects when used in consecutive test administrations; and
- (2) have a similar level of difficulty and thus diagnostic power to Reitan's TMT.

To achieve (1), the proposed approach uses pseudo-randomized processes to generate different instances. To achieve (2), our approach attempts to reproduce the spatial characteristics of Reitan's TMT to the greatest extent possible. According to [Vickers et al. \(1996\)](#), trails in Reitan's TMT are self-avoiding and gradually unwind in a clockwise or anticlockwise direction. To reproduce these characteristics, our DAC approach generates sub-solutions in divided problem spaces and combines sub-solutions to form a complete solution. In the "divide" phase, the test region is divided into several nested and non-overlapping layers. Within each layer, a partial trail is generated with the desired spatial characteristics. Then, in the "combine" phase, the partial trails are joined together to form a complete trail while preserving the desired characteristics.

With the increasing penetration and improved usability of digital devices, there has been a number of attempts to administer TMTs using digital devices, e.g. smartphones, tablets and computers. Computerized tests can facilitate more standardized and accurate data capturing as well as support detailed analysis. Following pre-designed test generation algorithms, they can also implement unlimited instances of TMT using systematic approaches. With these advantages, computerized tests also open the possibility to self-assessment in home environment. Traditionally, cognitive assessments need to be administered by healthcare professionals. With the aid of well-designed computerized tests, the assessment processes can be crowdsourced to the patients themselves.

To evaluate our DAC approach, we created a test application, called *iTMT*, which implements the DAC approach to generate computerized TMTs. A pilot study was conducted using this application and involving ten participants with different levels of cognitive abilities. The pilot study results suggest that no significant difference exists between the computerized tests generated by our DAC approach and Reitan's TMT in terms of total segment length and visual interference, indicating that they had a similar level of difficulty. Moreover, *iTMT* also demonstrated stronger construct validity, higher test-retest reliability and significantly reduced practice effects than existing computerized tests. These preliminary results support the effectiveness of the DAC approach and indicate that *iTMT* is a promising tool in longitudinal cognitive assessment.

In the following sections, we first introduce the neuropsychological background of TMT. Then, we review the prior efforts made to adapt it for longitudinal assessment and to computerize it. To address the issues identified from prior works, we propose our DAC approach in [Section 4](#). In [Section 5](#), we describe how we computerized our DAC approach and created *iTMT*. The pilot study results are presented in [Section 6](#). Finally, the main findings are presented in [Section 7](#).

2. Neuropsychological background of Trail Making Test

The TMT is one of the most frequently used neuropsychological tests ([Butler et al., 1991](#)) due to its sensitivity, simplicity and ease of administration. The TMT was originally constructed to assess general intelligence in [Army Individual Test Battery \(1944\)](#). It was later adapted by [Reitan \(1955\)](#) and included in the Halstead-Reitan Neuropsychological Battery in its current form. It is often used as a diagnostic tool to detect cognitive

impairments due to brain damage, e.g. dementia, stroke and traumatic brain injuries (Ashendorf *et al.*, 2008; Chen *et al.*, 2015).

During test administration of Reitan's TMT, the subject is instructed to connect a set of dots as quickly as possible while maintaining accuracy. Reitan's version only has one test instance and uses fixed dot arrangements. It contains two parts. As shown in Figure 1, Part A involves connecting 25 numbered dots in increasing order. Part B involves connecting 25 dots labelled with numbers and letters, in the alternating sequence "1-A-2-B-3-C[. . .]". The test is meant to assess cognitive functions such as cognitive flexibility, executive functioning, mental processing speed, divided attention as well as visual scanning (Sanchez-Cubillo *et al.*, 2009). Part B is more difficult than Part A, possibly due to differences in symbolic complexity and spatial arrangement (Fossum *et al.*, 1992). While Part A only contains numerals, Part B involves two symbol systems, alphabets and numerals, making it a more difficult cognitive task. Besides, *segment length* and amount of *visual interference* are also factors affecting the difficulty of TMT. Part B has a longer total trail length and more visual interfering stimuli than Part A, making it more demanding in terms of motor speed and visual scanning (Gaudino *et al.*, 1995).

Reitan's TMT is performed using paper and pencil and is administered by a healthcare professional. *Time to completion* and *frequency of errors* are the most common metrics which are recorded and used to interpret TMT performance of both Parts A and B.

3. Related work

3.1 Adapting Trail Making Test for longitudinal assessment

Although Reitan's TMT is widely used for assessing cognitive abilities, its use in longitudinal assessment is constrained due to its high susceptibility to practice effects

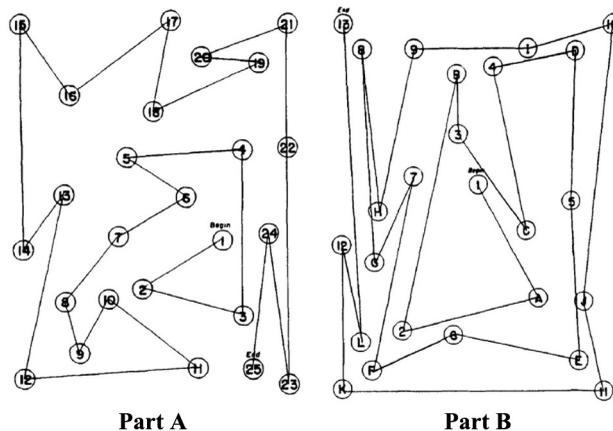


Figure 1.
The TMT: Parts A
and B

Notes: Part A (left) involves connecting 25 numbered dots in increasing order. Part B (right) involves connecting 25 dots labelled with numbers and letters, in the alternating sequence "1-A-2-B-3-C...".

(Beglinger *et al.*, 2005). Each repeated administration of it uses the same test instance, which decreases its sensitivity.

Atkinson and Ryan proposed to use other validated neuropsychological tests as alternative forms of TMT in longitudinal assessment (Atkinson and Ryan, 2007; Atkinson *et al.*, 2010). They identified three alternative forms that are equivalent and can be used interchangeably in a serial manner with a brief test–retest interval. However, due to their similarity in content and format, the use of alternative forms can only slightly reduce practice effects during longitudinal assessment. Instead of using other validated tests, Wagner *et al.* (2011) proposed to create new instances of TMT using its mirror images. Similar to the use of alternative forms, the mirror images also exhibited discernible practice effects when used serially in assessment.

The number of equivalent alternative tests and mirror images of TMT is limited. Hence, it is still not practical to adopt the aforementioned two approaches in longitudinal cognitive assessment over a long period. A group of researchers (Vickers *et al.*, 1996; Vickers and Lee, 1998) have proposed more systematic and divergent approaches for generating a theoretically unlimited number of TMT instances. They contended that, although it is not clear whether Reitan's TMT was constructed according to some implicit principles, it is vastly different from trails that are generated by purely random processes (Vickers *et al.*, 1996). They observed two characteristics of Reitan's TMT trails. First, for both Parts A and B, the trails gradually uncoil in either a clockwise or an anticlockwise direction from the inner to outer part. Second, the trails are self-avoiding, i.e. the line segments connecting consecutive points have no intersections with one another. However, it is a non-trivial problem to generate new instances of TMT which are endowed with these two characteristics (Vickers and Lee, 1998). Working reversely from the desired characteristics, two approaches to generate self-avoiding trails were proposed in Vickers *et al.* (1996) and Vickers and Lee (1998).

The first approach is suggested by considering the fractal nature of TMT trails. The problem of generating new TMT instance is transformed into the problem of generating self-avoiding fractal curve (Vickers *et al.*, 1996). Starting with a seed element, fractal curves are generated by repeatedly applying a set of transformations to the seed. The second approach transforms the problem of generating new TMT instance into a travelling salesman problem (TSP) (Vickers and Lee, 1998). Dots in TMT are treated as cities in TSP which can only be visited exactly once in a trail. The solution to TSP, the shortest path, is typically self-avoiding as having intersections in the trail tends to lengthen the path. Thus, given a set of dots, the solution to the corresponding TSP can be converted into a new instance of TMT.

Although both the fractal curve and the TSP approach can generate a theoretically unlimited number of TMT instances, the generated instances may have shorter average trail length and less visual interference than Reitan's paper-and-pencil TMT. Consequently, these instances may be less difficult than Reitan's version and less demanding in terms of the cognitive functions assessed. For example, less visual interfering stimuli would make the visual scanning much easier. Moreover, when test performance is measured by time to completion, the total segment length will also affect the test performance, as it affects the drawing time. Due to their reduced level of difficulty, the generated instances may not have enough discriminative power to produce statistically significant performance differences between different cognitive status. Consequently, their sensitivity and diagnostic efficacy may also be compromised.

3.2 Computerized Trail Making Test

Unlike manually administered tests which are susceptible to variations in test administrator and test procedures, computerized tests maintain standardized test procedures and are consistent across subjects (Woods *et al.*, 2015). Smith (2012) developed a computerized touch-screen version of TMT, eTrails, which contains a digital embodiment of the paper-and-pencil TMT and four computerized variants of the standard test. Experimental results suggest that all the five computerized tests have considerably higher test-retest reliability than TMT, possibly due to standardized procedures and less administrative errors. Another benefit brought by computerized tests is the ability to capture more high-fidelity data which increase the accuracy of performance measurements and facilitate in-depth performance analysis. Woods *et al.* (2015) developed a computerized TMT, C-TMT, which supports segment-by-segment analysis of performance and separates analysis of time spent on different tasks, e.g. dwelling and moving.

More importantly, computerized tests also provide the ideal format for generating a theoretically unlimited number of TMT instances. Following algorithmic approaches, pseudo-randomized dot arrangements can be generated for each test administration. Compared to Reitan's TMT, the practice effects for both eTrails and C-TMT were significantly attenuated (Smith, 2012; Woods *et al.*, 2015). Yet, it was not clearly indicated that how the instances of TMT were generated in these two computerized tests. The fractal curve approach and the TSP approach introduced in the previous section are both systematic and divergent approaches rooted in sound theoretical basis. However, to the best of our knowledge, neither of them has been formally computerized and validated in user studies.

We have reviewed the prior efforts made to adapt TMT for longitudinal assessment and to computerize TMT, as well as discussed the issues with these attempts. Next, we will propose our solution to the issues.

4. The divide-and-combine approach

In this section, we propose a DAC approach for generating different instances of TMT in a systematic manner. As introduced in Section 1, the generated instances need to be sufficiently different from each other while having a similar level of difficulty to Reitan's TMT. More specifically, the aim of the DAC approach is to produce different sets of ordered dots in a rectangular test region, so that the trail obtained by connecting dots in each set according to their order exhibits the following characteristics:

- uncoiling in clockwise or anticlockwise direction; and
- self-avoiding.

The DAC approach consists of two phases. In the “divide” phase, the test region is divided into several nested and non-overlapping layers. Within each layer, a partial trail is generated with the desired spatial characteristics. Then, in the “combine” phase, the partial trails are combined together to form a complete trail.

Suppose that we want to generate instances of TMT that contains n dots in a rectangular test region on the x - y plane. The plane spans across the area $x \in [a, b]$ and $y \in [c, d]$, where $a < b$ and $c < d$.

Suppose that we use m layers, where $m \in \mathbb{N}^+$. There are no constraints on the shape of the layers. However, the layers defined should satisfy two requirements. First, to make the generated trail follow an unwinding pattern, the layers need to be nested. Second, the layers should be non-overlapping to reduce intra-layer intersections.

Generally, the inner most layer can be a solid shape, while outer layers can be nested hollow shapes.

Definition 1: A layer L_i , where $i = 1, 2, \dots, m$, is defined as a 4-tuple $L_i = \{d_i, A_i, P_i, S_i\}$:

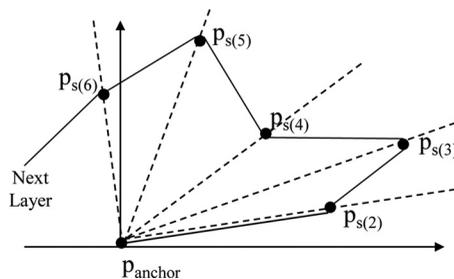
- $d_i \in \mathbb{R}_{>0}$ represents the relative density of the dots in layer L_i (refer to Definition 2). $d_i \times n$ gives the number of dots in layer L_i ;
- A_i defines the area layer L_i spans;
- $P_i = \{p_j = (x_j, y_j) | x_j \in [a, b]; y_j \in [c, d]; j = 1, 2, \dots, d_i \times n\}$ represents the dots in layer L_i ; and
- $S_i = \{s(k) | s(k) \in [1, d_i \times n]; k = 1, 2, \dots, d_i \times n\}$ represents the order of the dots in layer L_i . To form a partial trail T_i , the dots are connected according to the sequence of $p_{s(1)} \rightarrow p_{s(2)} \rightarrow \dots \rightarrow p_{s(d_i \times n)}$.

Definition 2: The density of dots D_i in a layer L_i can be defined as:

$$d_i = \frac{|P_i|}{n} \quad (0 < d_i \leq 1)$$

where d_i is a measure of relative density and is the ratio of number of dots in layer L_i (the cardinality of P_i) to the total number of dots n .

The pseudo code of the DAC approach is outlined in Algorithm 1. In the “divide” phase (Lines 3-5 of Algorithm 1), a partial trail T_i is generated within each layer. For each layer L_i , the dots are *randomly generated* within the layer and then sorted to determine their connecting order. The function *randomDots()* is called to generate $d_i \times n$ number of dots in the area defined by A_i . Then, the function *sortDots()* is invoked to sort the dots in layer L_i with respect to an anchor point and return their sorted order S_i . For the simplicity, the bottom left dot in the layer can be chosen as the anchor point in sorting. The dots are sorted according to their angles with respect to the anchor point. The order of sorting (increasing or decreasing angle) is determined randomly for each layer. A self-avoiding partial trail can be formed by traversing the dots by their sorted order. Figure 2 illustrates sorting by increasing angle.



Notes: The bottom left dot in the layer, P_anchor, is chosen as the anchor point in sorting. The dots are then connected according to the sorted order “ $s(1) - s(2) - \dots - s(6)$ ”, where $s(1)$ is the anchor point

Figure 2.
An illustration of sorting six dots in a layer by calling the function *sortDots()*

Algorithm 1: The DAC approach

```

1: var  $P_{regen}$                                 ▷ The dots that need to be regenerated
2: var  $p_{pivot}$                                 ▷ The pivot dot for combining two sub trails

3: for each  $L_i$  do                               ▷ The "divide" phase
4:    $P_i \leftarrow randomDots(A_i, d_i \times n)$ 
5:    $S_i \leftarrow sortDots(P_i)$ 

6: for  $i = 1$  to  $m - 1$  do                       ▷ The "combine" phase
7:   while  $intersect(T_i, T_{i+1})! = NULL$  do
8:      $P_{regen} \leftarrow intersect(T_i, T_{i+1})$     ▷ Dealing with intersections
9:      $P_{i+1} \leftarrow regenerateDots(P_{i+1}, P_{regen})$ 
10:     $S_{i+1} \leftarrow sortDots(P_{i+1})$ 
11:   for  $i = 1$  to  $m - 1$  do                       ▷ Connecting adjacent layers
12:      $p_{pivot} \leftarrow pickPivot(S_i, P_i, S_{i+1}, P_{i+1})$ 
13:      $S_{i+1} \leftarrow connect(S_i, S_{i+1}, p_{pivot})$ 
14:   for each  $P_i$  do
15:      $assignLabel(P_i, S_i)$ 

```

In the "combine" phase (Lines 6-15 of Algorithm 1), the partial trails in adjacent layers are combined together to form a complete trail. Before connecting the partial trails, a round of checking is performed to ensure that there are no intersections among them (Lines 6-10). The function $intersect(T_i, T_{i+1})$ is used to detect intersections between the partial trails in two adjacent layers, L_i and L_{i+1} . It returns the end points of line segments in the outer layer T_{i+1} that intersects with T_i . In the case when two trails intersect, i.e. $intersect(T_i, T_{i+1})! = NULL$, the trail in the outer layer will be adjusted to eliminate the intersection. The end points of the intersecting segments in T_{i+1} (or P_{regen}) are regenerated by calling the function $regenerateDots()$. With the newly generated dots, the dots in layer L_{i+1} are re-sorted by calling $sortDots()$ again. When no intersections are found, all the partial trails are connected together to form a complete trail (Lines 11-13). By calling $connect()$, each two partial trails in adjacent layers are joined through a pivot dot selected from the outer layer L_{i+1} . The function $pickPivot()$ is called to choose a pivot from P_{i+1} to connect T_i and T_{i+1} . The choice of pivot dot should guarantee that no intersection will be resulted from joining it and the last dot in the inner layer. Figure 3 illustrates connecting two trails L_{i-1} and L_i where $p_{s(3)}$ in L_i is chosen as the pivot dot. Finally, alphanumeric labels are assigned to the dots based on their order in the complete trail.

With the algorithm described above, the DAC approach is able to generate TMT instances that reproduce the two desired spatial characteristics: unwinding clockwise or counter-clockwise and self-avoiding. In the following part of this section, we will explain how the design of the DAC approach help to reproduce these characteristics.

4.1 Unwinding clockwise or counter-clockwise

The DAC approach uses both intra-layer and inter-layer designs to reproduce this characteristic. The intra-layer strategy ensures the partial trails exhibits the desired pattern locally within a layer, while the inter-layer strategy ensures that the pattern is embodied globally by a trail.

Within a layer (intra-layer), sorting and connecting the dots according to their angles with respect to the anchor point endows a partial trail with the desired unwinding pattern. As shown in Figure 2, the dots in a layer can also be viewed in a polar coordinate system with the anchor point as pole. Each dot can be represented by a radial coordinate (r, θ) ,

where $\theta \in [0, 2\pi]$. Connecting the dots according to the value of θ would form a trail that starts from and then zigzags around the anchor point. Visually, it can be viewed as a trail gradually uncoiling from the anchor point. The direction of unwinding can be controlled by the sorting order. If the angles are sorted in ascending order, the partial trail unwinds counter-clockwise. Conversely, the partial trail unwinds clockwise.

At the inter-layer level, the definition of the layers help to create a complete trail in the desired unwinding pattern. The layers defined must be nested. This requirement ensures that the area of an inner layer is surrounded by its outer layers. The trail formed by joining partial trails within nested layers also appears to be nested, which gradually unfolds when traversing from an inner layer to an outer layers.

4.2 Self-avoiding

To generate self-avoiding trails, the DAC approach also uses intra- and inter-layer strategies. The intra-layer strategy ensures each partial trail formed is self-avoiding. The inter-layer strategy generates a self-avoiding trail by connecting the partial trails while avoiding intersections in the connecting process.

At the intra-layer level, sorting and connecting dots by angles ensures that the resultant partial trail is self-avoiding. Referring to Figure 2 again, the area that spans across $[0, 2\pi]$ is divided into sectors by the dotted lines connecting the anchor point to other dots. These sectors are contiguous but non-overlapping. Each line segment of a partial trail spans within exactly one of the sectors. Thus, the line segments have no intersections with each other, except for the joints.

At the inter-layer level, using non-overlapping layers can help to reduce intersections among partial trails in different layers. However, a partial trail formed by traversing the dots within a layer may still cut through other layers. Intersection checking is performed to detect and eliminate intersections in such cases, as described by Algorithm 1 Lines 6 to 10. Moreover, the choice of pivot dot while connecting adjacent trails also avoids creating intersections in the combined trail.

The DAC approach is a generalized approach that is suitable for generating trails for both TMT Parts A and B. It maintains the modelling flexibility to accommodate the differences between Parts A and B. The dots in Part A scatter more evenly over the rectangular test region, while the dots in Part B are more skewed towards the rim of the

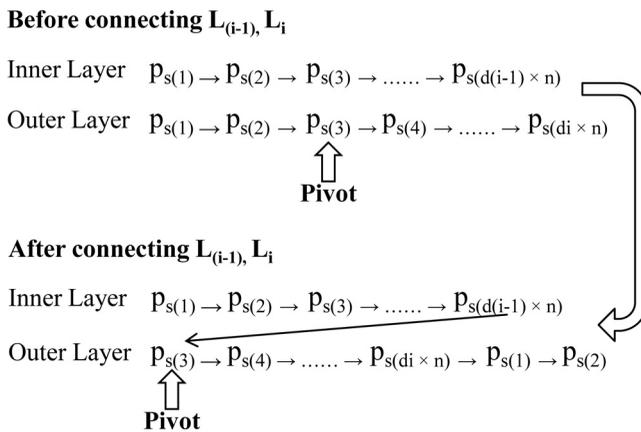


Figure 3. An illustration of connecting partial trails in two adjacent layers, $L_{(i-1)}$ and L_i . $P_{s(3)}$ in the outer layer (L_i) is chosen as the pivot point for connecting the two trails; the last point in the inner layer ($L_{(i-1)}$) is connected to $P_{s(3)}$

area. There are a number of strategies (Table I) to manipulate the distribution of the dots to cater to such differences.

In this section, we proposed the DAC approach for generating different instances of TMT. In the following sections, we will describe how this algorithmic approach can be implemented and built into computerized TMTs.

5. Computerizing Trail Making Test using the divide-and-combine approach

We created a test application called *iTMT*, which implements the DAC approach with the parameters determined from Reitan’s paper-and-pencil TMT.

After examining the dot arrangement in Reitan’s TMT, we chose to use 25 dots and three layers for both Part A and Part B, i.e. $n = 25, m = 3$. For the ease of modelling, we used three concentric layers, with each of their centres at the origin. As illustrated in Figure 4, L_1 is a rectangle, while L_2 and L_3 are hollow rectangles. Collectively, the three non-overlapping layers cover the test region exhaustively. The values of d_i are presented in Table II. Compared to Part A, Part B has smaller relative density in L_1 and L_2 , but much greater density in L_3 , which is designed in accordance with the observations made from the paper-and-pencil TMT.

Table I.
Strategies for generating TMT Part A and Part B with the DAC approach

Parameter to adjust	Part A	Part B
Relative density of dots in a layer, d_i	Assign larger d_i to inner layers, smaller d_i to outer layers	Assign smaller d_i to inner layers, larger d_i to outer layers
Area of a layer, A_i	Define smaller A_i for inner layers, larger A_i for outer layers	Define larger A_i for inner layers, smaller A_i for outer layers
Total number of layers, m	Use larger m for Part A	Use smaller m

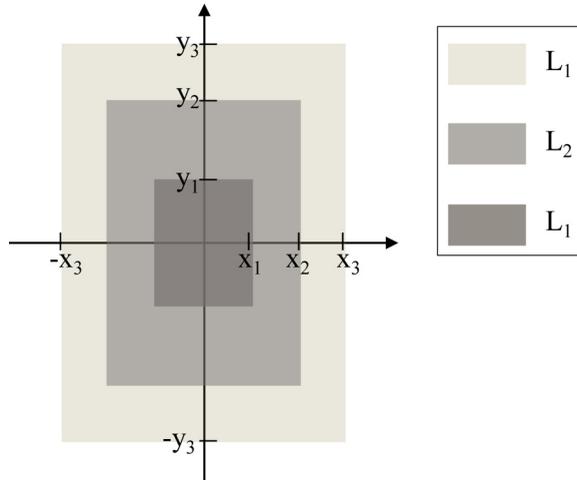


Figure 4.
The three layers used in the implementation of *iTMT* ($m = 3$)

Notes: L_1 is a rectangle, while L_2 and L_3 are hollow rectangles. Collectively, the three non-overlapping layers cover the test region exhaustively

The area each layer spans can be defined as following:

$$A_i \text{ spans the area } \begin{cases} |x| \leq x_1, |y| \leq y_1 & i = 1 \\ x_1 < |x| \leq x_2, y_1 < |y| \leq y_2 & i = 2 \\ x_2 < |x| \leq x_3, y_2 < |y| \leq y_3 & i = 3 \end{cases}$$

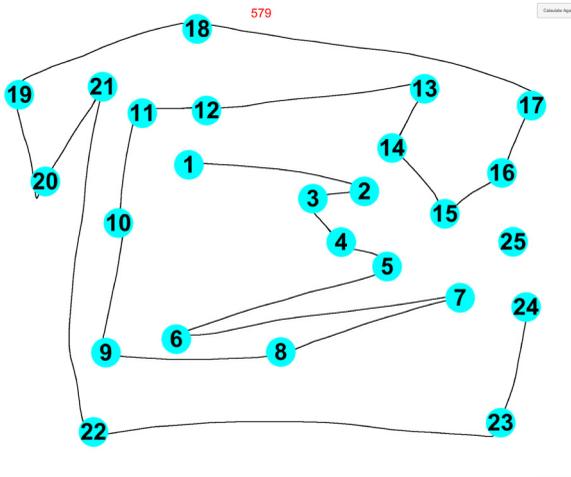
Following the strategies in Table I, A_i is defined differently for Parts A and B. Comparatively, for Part B, A_1 spans across larger area, while A_2 and A_3 are made smaller so that more dots would be placed near the edge of the test region.

Using the aforementioned parameters, we built the test application, *iTMT*. It is a touch-screen application that can administer computerized tests on tablets. Instead of using a pencil, the subjects need to connect the dots using their fingers. During each administration, *iTMT* generates a new instance of TMT using the DAC approach. Figure 5 shows a test screenshot of *iTMT*, where Part A was being administered.

6. Evaluating computerized Trail Making Test

To evaluate the DAC approach, a pilot study was conducted for *iTMT*. The instances of computerized TMT generated with the DAC approach was evaluated for the following aspects:

Density	d_1	d_2	d_3	Table II. Relative densities d_i for each layer in Part A and Part B
Part A	0.24	0.36	0.4	
Part B	0.2	0.2	0.6	



Note: The trail formed by connecting the dots in sequence exhibits the two desired characteristics: uncoiling in clockwise or anticlockwise direction, and self-avoiding

Figure 5.
A screenshot of *iTMT* Part A, correctly connected by a test subject

- *Level of difficulty*: Whether the generated instances are equivalent to Reitan's paper-and-pencil TMT in terms of test difficulty (as measured by total segment length and visual interference).
- *Construct validity*: Whether the generated instances are able to produce test results that are comparable to Reitan's paper-and-pencil TMT (as measured by time to completion).
- *Test-retest reliability*: Whether the test results are consistent across different instances (as measured by time to completion) and whether there are significant practice effects.

The study involved ten participants aged from 33 to 84 years, with 8 above 50 and 4 above 65. As age is an important factor which affects cognitive abilities, we recruited participants from a broad age range to include people with different levels of cognitive abilities. Considering that some elderly participants are illiterate or only had limited education, *only Part A* of TMT was administered. Each subject was administered three tests consecutively, including one paper-and-pencil test and two computerized tests on a tablet. The computerized tests were generated by *iTMT* using the DAC approach. As the size of test regions may affect the difficulty metrics, we used an adapted version of Part A in Reitan's TMT for the paper-and-pencil test to control this factor. Reitan's Part A is compressed proportionally into a smaller test region which has an equal area as the screen of the tablet used, with the relative positions of the dots preserved.

Before conducting the tests, the participant was shown a sample TMT with eight dots so as to become familiar with the connecting rules and test procedures. Then, the paper-and-pencil test was administered to the participant. The time taken to complete the trail, i.e. time to completion, was recorded with a timer. After an interval of 30 s, two computerized tests were administered to the participant consecutively. The test statistics were recorded by *iTMT* for later analysis.

6.1 Level of difficulty

As proposed in Gaudino *et al.* (1995), we use total segment length and visual interference as two metrics for the difficulty of TMTs. Total segment length is defined as the summed length of the shortest line segments connecting successive dots. Visual interference is quantified by summing the number of dots that lie within 3 cm of each line segment. The two metrics of the paper-and-pencil TMT administered in the study were calculated and used as normative values for evaluating the computerized tests. The paper-and-pencil TMT has a total segment length of 135.9 cm and visual interference of 49. For each of the 20 computerized tests (2 per participant \times 10) administered during the study, its total segment length and visual interference was calculated. Two *t*-tests were performed to determine whether the computerized tests generated by *iTMT* have a similar level of difficulty to the paper-and-pencil test (Table III). One *t*-test compared the mean of total segment length (μ_{seg}) of computerized tests and the total segment length of the paper-and-pencil test with the null hypothesis $\mu_{seg} = 135.9$. As $p = 0.33 > \alpha = 0.05$ ($t_0 = 0.98 < t_{0.05,19} = 2.093$), we accept the null hypothesis. The other *t*-test compared the mean of visual interference (μ_{vis}) of computerized tests and the visual interference of the paper-and-pencil test with the null hypothesis $\mu_{vis} = 49$. As $p = 0.31 > \alpha = 0.05$ ($t_0 = 1.04 < t_{0.05,19} = 2.093$), we also accept H_0 .

The test results suggest that there is no significant difference between the computerized tests and the paper-and-pencil test in terms of total segment length and visual interference.

Hence, the results further indicate that, as measured by these two metrics, computerized tests and the paper-and-pencil test can be considered to have a similar level of difficulty.

6.2 Construct validity

The construct validity of the computerized tests generated by *iTMT* was measured by correlating the time taken by the same participant to complete a computerized test and the paper-and-pencil test. The Pearson product-moment correlations between time to completion are shown in Figure 6. The time taken by a participant to complete the first computerized test was positively correlated with the time he/she needs to complete the paper-and-pencil test ($r = 0.89, p = 0.0006$). Positive correlation was also reported between the time to completion of the second computerized test and the paper-and-pencil test, with a higher correlation coefficient ($r = 0.97, p = 0$). Both correlations reached statistical significance ($p < 0.01$), indicating significant linear relationships between the time to completion of computerized tests and paper-and-pencil test. Compared to eTrails (Smith, 2012) (Table IV), an existing computerized test that achieved moderate correlation with the paper-and-pencil test (highest $r = 0.668$), *iTMT* demonstrated a much stronger correlation with the paper-and-pencil test ($r = 0.89$ and 0.97). When test performance was measured by time to completion, the computerized tests generated by *iTMT* were able to produce test scores that are highly correlated with the paper-and-pencil test score of the same participant, suggesting high construct validity of the computerized tests.

H_0	\bar{x}	p -value	t score	t_0	Result
Segment length $\mu_{seg} = 135.9$	141	0.33	0.98	2.093	Accept H_0
Visual interference $\mu_{vis} = 49$	51.19	0.31	1.04	2.093	Accept H_0

Table III. Hypothesis tests for segment length and visual interference ($\alpha = 0.05$)

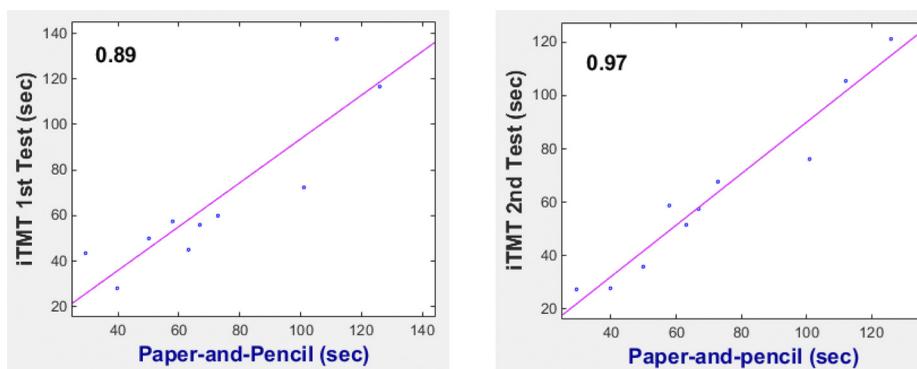


Figure 6. The correlation between time to completion of computerized tests and paper-and-pencil test

Notes: Left: positive correlation between first *iTMT* test and paper-and-pencil test ($r = 0.89, p = 0.0006$). Right: positive correlation between second *iTMT* test and paper-and-pencil test ($r = 0.97, p = 0$)

6.3 Test-retest reliability

The test-retest reliability of the computerized tests generated by *iTMT* was calculated by correlating the time taken to complete the first and the second computerized test. The Pearson product-moment correlation between the time to completion of two computerized tests is shown in Figure 7. Significant linear relationship was observed between the time to completion of the two computerized tests ($r = 0.93$). *iTMT* achieved considerably higher test-retest reliability than eTrails (highest $r = 0.62$) (Smith, 2012) and C-TMT (ICC = 0.87) (Woods *et al.*, 2015). Moreover, *iTMT* was found to have *almost no discernible practice effects*. As shown in Figure 7, some participants took longer to complete the first test, while others took longer to complete the second test on *iTMT*. From the first to the second administration, the average time to completion was only reduced by 5.58 per cent. While the reduction in average time to completion for eTrail (Smith, 2012) was 10.33 and 12 per cent for C-TMT (Woods *et al.*, 2015) (Table IV). Comparatively, *iTMT* was less susceptible to practice effects, providing further evidence to support its high reliability when administered repeatedly.

7. Conclusion

In this paper, we proposed a DAC approach to generate instances of TMT that can be used in longitudinal cognitive assessment. Our proposed approach is able to generated a theoretically unlimited number of different TMT instances which can be used in consecutive test administrations. Moreover, the instances generated by the proposed approach have a similar

Table IV.
Comparing the validity and reliability of *iTMT* and other existing computerized TMTs

Name	Construct validity (measured by r)	Test-retest reliability (measured by r)	Test-retest reliability (measured by % reduction)
<i>iTMT</i>	0.89, 0.97	0.93	5.58
eTrails	0.668	0.62	10.33
C-TMT	–	0.87	12

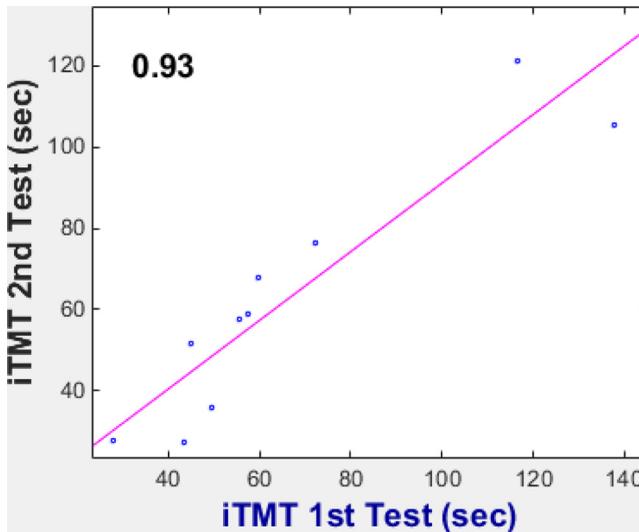


Figure 7.
The correlation between time to completion of the first *iTMT* test and the second *iTMT* test ($r = 0.93$)

level of difficulty to Reitan's paper-and-pencil TMT by reproducing its spatial characteristics. We also created a test application, *iTMT*, which implements the DAC approach to generate computerized TMTs. The preliminary results from the pilot study support the effectiveness of our DAC approach. Compared to existing computerized tests, the instances of TMT generated by the approach produced test results that were significantly more correlated with results of Reitan's TMT. Similar difficulty and highly correlated results suggest that these instances possess similar diagnostic power to the Reitan's TMT and are able to better distinguish subjects with different level of cognitive abilities. *iTMT* also demonstrated higher test-retest reliability and significantly reduced practice effects than existing computerized tests. Due to the illiteracy of some participants, only TMT Part A and its computerized versions were tested in this study. In the future, we plan to study Part B in a similar way.

By supporting self-assessment, *iTMT* also can help to crowdsource the assessment processes, which need to be administered by healthcare professionals conventionally, to the patients themselves. To validate the feasibility of using *iTMT* in self-assessment over a long time period, it is worthwhile to study how test performance of it would change beyond the 2nd administration. Further experiments need to be conducted to find out whether the test performance of *iTMT* is convergent and how does the test performance converge.

Note

1. Data retrieved from *Dementia Statistic* website by Alzheimer's Disease International, accessible till December 2016, available at: www.alz.co.uk/research/statistics

References

- Army Individual Test Battery (1944), *Manual of Directions and Scoring*, War Department, Adjutant Generals Office, Washington, DC.
- Ashendorf, L., Jefferson, A.L., O'Connor, M.K., Chaisson, C., Green, R.C. and Stern, R.A. (2008), "Trail Making Test errors in normal aging, mild cognitive impairment, and dementia", *Archives of Clinical Neuropsychology*, Vol. 23 No. 2, pp. 129-137.
- Atkinson, T.M. and Ryan, J.P. (2007), "The use of variants of the Trail Making Test in serial assessment: a construct validity study", *Journal of Psychoeducational Assessment*, Vol. 26 No. 1, pp. 42-53.
- Atkinson, T.M., Ryan, J.P., Lent, A., Wallis, A., Schachter, H. and Coder, R. (2010), "Three Trail Making Tests for use in neuropsychological assessments with brief intertest intervals", *Journal of Clinical and Experimental Neuropsychology*, Vol. 32 No. 2, pp. 151-158.
- Beglinger, L.J., Gaydos, B., Tangphao-Daniels, O., Duff, K., Kareken, D.A., Crawford, J., Fastenau, P.S. and Siemers, E.R. (2005), "Practice effects and the use of alternate forms in serial neuropsychological testing", *Archives of Clinical Neuropsychology*, Vol. 20 No. 4, pp. 517-529.
- Buck, K.K., Atkinson, T.M. and Ryan, J.P. (2008), "Evidence of practice effects in variants of the Trail Making Test during serial assessment", *Journal of Clinical and Experimental Neuropsychology*, Vol. 30 No. 3, pp. 312-318.
- Butler, M., Retzlaff, P.D. and Vanderploeg, R. (1991), "Neuropsychological test usage", *Professional Psychology Research and Practice*, Vol. 22 No. 6, pp. 510-512.
- Chen, Y., Yu, H., Miao, C., Chen, B., Yang, X. and Leung, C. (2015), "Using motor patterns for stroke detection", *Science (Advances in Computational Psychophysiology)*, Vol. 350 No. 6256, pp. 12-14.
- Fossum, B., Holmberg, H. and Reinvang, I. (1992), "Spatial and symbolic factors in performance on the Trail Making Test", *Neuropsychology*, Vol. 6 No. 1, pp. 71-75.

- Gaudino, E.A., Geisler, M.W. and Squires, N.K. (1995), "Construct validity in the Trail Making Test: what makes part B harder?", *Journal of Clinical and Experimental Neuropsychology*, Vol. 17 No. 4, pp. 529-535.
- Lyketsos, C.G., Lopez, O., Jones, B., Fitzpatrick, A.L., Breitner, J. and DeKosky, S. (2002), "Prevalence of neuropsychiatric symptoms in dementia and mild cognitive impairment: results from the cardiovascular health study", *JAMA*, Vol. 288 No. 12, pp. 1475-1483.
- Muir, S.W., Gopaul, K. and Odasso, M.M.M. (2012), "The role of cognitive impairment in fall risk among older adults: a systematic review and meta-analysis", *Age Ageing*, Vol. 41 No. 3, pp. 299-308.
- Reitan, R.M. (1955), "The relation of the Trail Making Test to organic brain damage", *Journal of Consulting and Clinical Psychology*, Vol. 19 No. 5, pp. 393-394.
- Sanchez-Cubillo, I., Periáñez, J.A., Adrover-Roig, D., Rodríguez-Sánchez, J.M., Ríos-Lago, M., Tirapu, J. and Barceló, F. (2009), "Construct validity of the Trail Making Test: role of task-switching, working memory, inhibition/interference control, and visuomotor abilities", *International Neuropsychological Society*, Vol. 15 No. 3, pp. 438-450.
- Sattin, R.W. (1992), "Falls among older persons: a public health perspective", *Annual Review of Public Health*, Vol. 13 No. 1, pp. 489-508.
- Smith, B.T. (2012), "Creation of a more accurate and predictive Trail Making Test", Doctoral dissertation, University of North Carolina Wilmington.
- Tinetti, M.E., Speechley, M. and Ginter, S.F. (1988), "Risk factors for falls among elderly persons living in the community", *New England Journal of Medicine*, Vol. 319 No. 26, pp. 1701-1707.
- Vickers, D. and Lee, M.D. (1998), "Never cross the path of a traveling salesman: the neural network generation of halstead-reitan Trail Making Tests", *Behavior Research Methods, Instruments, & Computers*, Vol. 30 No. 3, pp. 423-431.
- Vickers, D., Vincent, N. and Medvedev, A. (1996), "The geometric structure, construction, and interpretation of pathfollowing (trailmaking) tests", *The Journal of Clinical Psychology*, Vol. 52 No. 6, pp. 651-661.
- Wagner, S., Helmreich, I., Dahmen, N., Lieb, K. and Tadic, A. (2011), "Reliability of three alternate forms of the Trail Making Tests a and B", *Archives of Clinical Neuropsychology*, Vol. 26 No. 4, pp. 314-321.
- Woods, D.L., Wyma, J.M., Herron, T.J. and Yund, E.W. (2015), "The effects of aging, malingering, and traumatic brain injury on computerized trail-making test performance", *PLoS one*, Vol. 10 No. 6, p. e0124345.

Further reading

- Drapeau, C.E., Bastien-Toniazzo, M., Rous, C. and Carlier, M. (2007), "Nonequivalence of computerized and paper-and-pencil versions of Trail Making Test", *Perceptual and Motor Skills*, Vol. 104 No. 3, pp. 785-791.
- Fryer, S., Sutton, E., Tiplady, B. and Wight, P. (2000), "Trail making without trails: the use of a pen computer task for assessing effects of brain injury", *Clinical Neuropsychological Assessment*, Vol. 2, pp. 151-165.
- Tombaugh, T.N. (2004), "Trail Making Test a and b: normative data stratified by age and education", *Archives of Clinical Neuropsychology*, Vol. 19 No. 2, pp. 203-214.

About the authors

Zhiwei Zeng received the BEng (Hons) degree in Computer Science and BBus (Hons) degree from Nanyang Technological University (NTU) in 2015, where she is currently pursuing a PhD degree at the NTU-UBC Joint Research Centre of Excellence in Active Living for the Elderly (LILY), Interdisciplinary Graduate School. Her current research interests include artificial persuasion, its

computational modelling and applications of intelligent agents in healthcare systems. Zhiwei Zeng is the corresponding author and can be contacted at: il60001@e.ntu.edu.sg

Chunyan Miao is a Professor with the School of Computer Science and Engineering (SCSE) at Nanyang Technological University (NTU), Singapore. She is the Director of the NTU-UBC Joint Research Centre of Excellence in Active Living for the Elderly (LILY). Prior to joining NTU, she was an Instructor and a Post-Doctoral Fellow with the School of Computing, Simon Fraser University, Canada. Her research focuses on studying the cognitive and social characteristics of intelligent agents in multi-agent and distributed AI/CI systems, such as trust, emotions, incentives, motivated learning, ecological and organizational behavior. She has worked on new disruptive Artificial intelligence (AI) approaches and theories that synergize human intelligence, artificial intelligence and behavior data analytics (AI powered by humans). Her current research interests include human-agent interaction, cognitive agents, human computation and serious games.

Cyril Leung is a Member of the IEEE and the IEEE Computer Society. He received the BSc (Hons) degree from Imperial College, University of London, England, UK, in 1973, and the MS and PhD degrees in electrical engineering from Stanford University, Stanford, CA, USA, in 1974 and 1976, respectively. From 1976 to 1979, he was an Assistant Professor with the Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, MA, USA. From 1979 to 1980, he was with the Department of Systems Engineering and Computing Science, Carleton University, Ottawa, ON, Canada. Since July 1980, he has been with the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC, Canada, where he is a Professor and currently holds the PMC-Sierra Professorship in Networking and Communications. He is the Deputy Director of the NTU-UBC Joint Research Centre of Excellence in Active Living for the Elderly (LILY). His current research interests include wireless communications systems. He is a Member of the Association of Professional Engineers and Geoscientists of British Columbia, Canada.

Zhiqi Shen is currently with the School of Computer Science and Engineering (SCSE), Nanyang Technological University, Singapore. He received the BSc in Computer Science and Technology from Peking University, Beijing, China, the MEng in Computer Engineering with the Beijing University of Technology, Beijing, and the PhD with Nanyang Technological University, Singapore. His current research interests include artificial intelligence, software agents, multi-agent systems (MAS); goal oriented modelling, agent oriented software engineering; internet of things, crowdsourcing, e-learning, agent augmented interactive media, game design and interactive story telling.

For instructions on how to order reprints of this article, please visit our website:

www.emeraldgrouppublishing.com/licensing/reprints.htm

Or contact us for further details: permissions@emeraldinsight.com