Djehuty: A Mixed-Initiative Handwriting Game for Preschoolers

Jean Michel A. Sarr*
Université Cheikh Anta Diop de Dakar (UCAD)
Dakar, Senegal
jm.amath@icloud.com

Georgios N. Yannakakis
Institute of Digital Games
University of Malta
Msida, Malta
georgios.yannakakis@um.edu.mt

Antonios Liapis
Institute of Digital Games
University of Malta
Msida, Malta
antonios.liapis@um.edu.mt

Alassane Bah†
Université Cheikh Anta Diop de Dakar (UCAD)
Dakar, Senegal
alassane.bah@gmail.com

Christophe Cambier
Sorbonne Université, IRD, UMMISCO
Paris, France
christophe.cambier@ird.fr

ABSTRACT

Learning to read and write is a fundamental right and a necessary skill for the personal, cultural, and economic development of people and their societies. However, children of developing countries, such as sub-Saharan areas, are currently at a greater risk of illiteracy. The current penetration of mobile technologies and the internet in sub-Saharan rural areas, however, offers a unique opportunity for tackling the challenge of literacy at a large scale. Motivated by the current shortage of preschool teachers for training handwriting in a personalised manner, this paper discusses the design of Djehuty, an educational gamified environment for preschoolers. Djehuty is equipped with an artificial intelligence module which generates a style of handwriting and suggests handwriting paths to the child in a *mixed-initiative* manner. The paper presents the key elements of the game prototype.

CCS CONCEPTS

• **Applied computing** → **Computer-assisted instruction**; *Computer games*; • **Computing methodologies** → *Neural networks*; • **Human-centered computing** → *Touch screens.*

KEYWORDS

intelligent tutoring system, mixed-initiative generation, early childhood education, generative models, handwriting instruction

ACM Reference Format:
Jean Michel A. Sarr, Georgios N. Yannakakis, Antonios Liapis, Alassane Bah, and Christophe Cambier. 2020. Djehuty: A Mixed-Initiative Handwriting Game for Preschoolers. In *International Conference on the Foundations of Digital Games (FDG '20), September 15–18, 2020, Bugibba, Malta.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3402942.3403030

*Also with Sorbonne Université, IRD, UMMISCO.
†Also with Ecole Supérieure Polytechnique.

FDG '20, September 15–18, 2020, Bugibba, Malta

ACM ISBN 978-1-4503-8807-8/20/09…$15.00
https://doi.org/10.1145/3402942.3403030

1 INTRODUCTION

According to the World Bank [4], sub-Saharan countries have the highest population growth rate in the world (currently 2.7% annually). Only one out of four children in these countries, however, is likely to enroll into preschool [29]. As a result, the literacy rate in sub-Saharan Africa is currently the lowest in the world [3]. Indicatively, in 2018 only 64% of adults could read and write. Such figures demonstrate that the *Sustainable Development Goal 4* [28]—which envisages quality education for all—remains a grand challenge in sub-Saharan Africa. As smartphone usage in these areas is growing fast [25] thanks to low-cost entry-levels, we believe that designing a game for mobile devices will have a comparably larger adoption and direct benefits for improving the writing and reading skills of these populations. We argue that if the game can assist the handwriting process, learning to write may happen in a personalized manner. Motivated by the lack of well trained preschool and elementary teachers for teaching handwriting to children in rural areas, this paper introduces *Djehuty*, a gamified Intelligent Tutoring System (ITS) that features a mixed-initiative letter generator that assists preschoolers in learning to write. The artificial intelligence (AI) system operates in an *iterative refinement* fashion [20], observing the way children write a letter and then proposing a new letter which is closer to the typical way this letter is written. The paper describes the key elements of the game prototype.

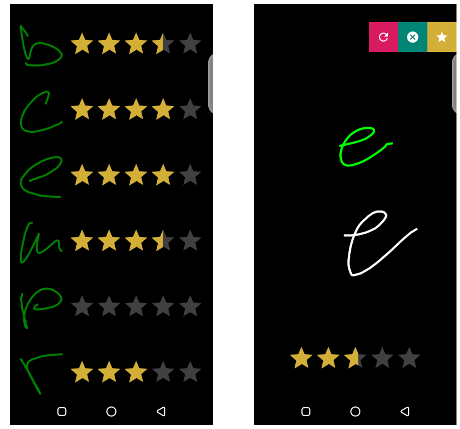
2 RELATED WORK

Artificial intelligence has been applied for the purposes of education for 50 years, with intelligent tutors such as Jamie Carbonell's SCHOLAR [7]. One of the core goals of AI in education is "to match the needs of individual students by providing alternative representations of content, alternative paths through material, and alternative means of interaction" [31]. Mixed-initiative dialogue between learner and AI has often been employed for the purposes of education, including in the case of SCHOLAR. While the exact nature of *mixed-initiative* interaction has not been formally defined [24], we follow the premise of [32] that "both the human and the computer proactively make contributions to the problem solution, although the two initiatives do not need to contribute to the same degree". Intelligent tutoring systems are the most common applications of AI in education. In her seminal book on ITS [31], Woolf proposed seven features that a system might have to be considered as an ITS: (1) *Student modeling*, i.e. making an accurate representation of student knowledge, to adapt to student preferred learning style and enhance learning; (2) *Expert modeling*, i.e. the representation of the knowledge to teach (e.g. geography, algebra, etc.); (3) *Generativity*, i.e. the ability for a system to generate problems for the student, or given a problem to generate appropriate tips to help her learn a concept; (4) *Mixed-initiative*, i.e. the ability for either the student or the tutor to take control of the interaction to achieve learning benefits or more engagement; (5) *Instructional modeling*, i.e. how a tutor modifies its guidance to improve learning outcomes for each student; (6) *Self-improvement*, i.e. the ability of a tutor to modify its performance based on its experience with prior students; (7) *Interactive learning* i.e. a system that is responsive to the student.

2.1 ITS for handwriting

Handwriting teaching technologies have been developed for multiple writing systems: the Latin alphabet [1, 10, 15, 27], Arabic scripts [5, 12], Bengali alphabet [18], as well as Kanji characters [2, 16, 17]. In this review, we classify ITS in levels based on the number of features they implement: e.g. an ITS is said to be level 4 if it has four ITS features. Only works that implemented at least 4 components of ITS in the context of teaching handwriting were surveyed. In particular, every system presented below implements *interactivity*, *interactive learning* and *student modeling*, and the following sections classify systems by level and highlight their differentiating factors.

2.1.1 Level 4 with expert modeling. In 2002, Ando *et al.* [2] proposed a system to learn Kanji writing. The student can draw a Kanji with the computer mouse, which is compared with the reference model of 2000 classic characters before a correction is issued. Kanjis are represented using the three-point approximation that corresponds to these characters. To evaluate the student, the system compares the order of lines produced with the reference. The comparison between student-drawn and template characters was refined by Hu *et al.* [16, 17], who represented Kanji with an Attributed Relational Graph (ARG). This approach could automatically detect errors in the character plot by associating the student's sample with the ground truth directly at the graph level. The system associates and generates an error description, if necessary. In 2010, Trazo [10] was introduced for teaching cursive writing to children of pre-school age (3-4 years) on a tablet. The system has two modules, one for the child and one for the teacher (for administering and monitoring the students). The child can exercise in drawing horizontal lines, vertical lines or simple curves. The system records the child's path and measures the stroke size and speed when drawing. Trazo is also able to compare the child's drawing with a pre-determined model. The first ITS for Arabic calligraphy was implemented in [5] and designed for tablets. It encodes several models through an ARG as in [17], since it can represent the relationships between the many plots of a pattern. In 2016, the Arab Kit Tutor [12] was introduced as an Android application with an interface for both students and teachers, using a multi-agent approach to analyze the production of students. In 2016, a web application was introduced for teaching Bengali characters [18]. Several examples of correctly written characters are stored in memory on the server and constitute the expert model. The student has the opportunity to follow a guided learning mode or free learning.



(a) (b)

Figure 1: Screenshots of learners' interaction with Djehuty. (a) Letter selection, showing the learner's highest score per letter. (b) The computer generates and animates a letter sample (in green), the user draws (in white) and presses the yellow "star" button to receive a score in real-time.

2.1.2 Level 4 with generativity. Hood *et al.* reversed the usual pattern of expert feedback on student's output via a social robot which heeds student's advice for improving [15]. The interaction with the system is as follows: (a) children give a card with a word to write to the robot, (b) the robot writes the word seen and asks for feedback, (c) the child provides feedback to the robot via a demonstration, (d) the robot retries until the user is satisfied. This ITS has an expert model (a statistical representation of the letters' shape), while it can also generate handwritten letters and interact with the child.

2.1.3 Level 5 with expert modeling & instructional modeling. The handwriting ITS of Albu *et al.* adapted feedback on the quality of a student's production according to their affective state[1]. While the ITS is used the student's emotions are classified via their voice and facial data. A pre-test questionnaire also assesses the student's personality. Finally, the quality of letters is compared to the expert model via Dynamic Time Warping, which computes the similarity between time series. This allows the tutor to select a pedagogical strategy adapted to the child's affective state and skills. Rather than focus on the student's emotions, Simonnet *et al.* [27] attempt to adjust the difficulty of the task. In their tablet-based digital notebook for handwriting, the student has two types of models to reproduce: a dynamic model (i.e an animation) or an image. The ITS provides a personalized correction based on the shape of the letter, the direction, the order of plots, or a combination of the above. Based on this analysis, the system adapts to the level of each learner by providing exercises of an adequate level of difficulty.

3 DJEHUTY

Djehuty (better known by his Greek name, Thoth) was the god of wisdom, writing, hieroglyphics, science and the arts for the ancient Egyptians. This is the name we have chosen for our interactive tutoring system to teach Latin alphabet handwriting to preschoolers. Djehuty is a gamified ITS and it implements three features of gamified systems [22]: simplicity, feedback and real-time. It is also a level 5 ITS with a mixed-initiative and a generative component.

3.1 Gamification features

The player's objective in Djehuty is simple: to reproduce the letter drawn on the screen as accurately as possible. See Figure 1 for an indicative playthrough with Djehuty. The steps are: (1) The player requests the application to write a letter on the screen. The player can sample many examples of the same letter in real-time, by clicking on the refresh button. (2) The player tries to replicate the letter she wants, by drawing on the screen. (3) By clicking on the star button, the student gets a score in stars. The more accurate the replication, the higher the score. If the user got more than 3.5 stars, the the next letter is sampled in the style of the user. (4) The player can retry as many times as she wants, as the AI produces a new shape for the same letter after a player gets a score; otherwise, the player can exit to the main menu with the back button. The goal is to get as many stars as possible in all letters.

3.2 ITS features

The game content can be likened to the expert model. For Djehuty, a variational autoencoder (VAE) [19] was trained on an open-source dataset of letter drawing [21] to build the expert model. Eight letters are implemented: a, b, c, e, m, p, r, s (see Fig. 1a).

3.2.1 Expert model and generativity. A sketch-RNN [14] is a sequence to sequence VAE that uses a recurrent neural network (RNN) as the main architecture. The sketch-RNN was designed to learn sequences of sketches corresponding to people's drawings such as birds, fruits, trucks, etc. Djehuty uses a simplified version of sketch-RNN, with two minor modifications. Gated Recurrent Unit (GRUs) [8] are used instead of LSTMs to reduce overall model complexity. GRUs need fewer parameters but are competitive compared to LSTM [8]. Secondly, Djehuty only uses diagonal covariance matrices instead of using full covariance matrices within the mixture of Gaussian functions. The paper will still refer to the simplified model described above as Sketch-RNN, as it only differs slightly from the original implementation. After training, we exported the sketch-RNN and also the standalone decoder with a latent vector for each letter, in order to allow multiple scenarios to run efficiently on the Android system. Generativity provides a desirable difficulty [6] to learners due to the infinite variations of letters produced when noise is added to the sketch-RNN's decoder. Studies have shown the developmental character of writing [13], as 3-year old children produce undifferentiated forms, scribbling, wavy lines, and pictorial representations. Thus we interpret the notion of writing a letter from the standpoint of a child as a creative process involving lateral thinking. Lateral thinking [9] is the process of solving seemingly unsolvable problems or tackling non-trivial tasks through an indirect, non-linear, creative approach. Random stimuli is the main guarantor of foreignness to stimulate creativity. Indeed, each letter requested by the player is different. The purpose is to disrupt preconceived notions and habitual patterns of learning, by forcing the user to integrate and/or exploit each new element in the creation of its own representation of the letter. By re-framing [32] the way she writes in order to emulate the newly generated variation of the letter, the learner becomes more adept while also not perceiving the task as repetitive.

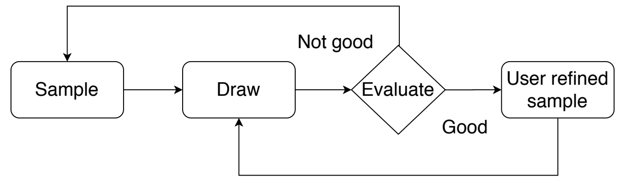


Figure 2: Mixed-initiative game loop: The user first samples a letter (generated by sketch-RNN) and draws it on the screen. Her drawing is evaluated based on similarity to the AI-drawn letter: if the result is less than 3.5 stars then a letter is sampled again (using only the latent space and the decoder). Otherwise the user's letter forms the input of sketch-RNN which produces a version of the letter in the style of the user, following an iterative refinement approach.

3.2.2 Student model. We evaluate the angular distance between the student trial and the sample to get a measure of performance. The angular similarity derives from the cosine similarity between two vectors. Metrics based on cosine similarity have been used to compare two objects of the same dimension. Examples range from speaker recognition [11] to face verification [23].

3.2.3 Mixed-initiative. To a degree, the fact that the computer proactively generates a new challenge for the learner in the form of a variation of a letter makes Djehuty a mixed-initiative system. While the learner chooses which letter to draw, leading the task initiative [24], the AI proactively refines the task by posing the specific challenge. More importantly, however, Djehuty is a mixed-initiative system as it accounts for the learner's input when creating new letters for her to draw. After the first attempt of a learner to draw a letter, if her score was 3.5 stars or above then sketch-RNN uses the learner's character as input: this means that the AI-drawn character will be a variation of the learner's previous character. As mentioned in Section 3.2.1, two neural network models are deployed in the Android system for each letter: the sketch-RNN and the standalone decoder along with a latent vector. When the user samples another letter variant with the refresh button, Gaussian noise is injected to the latent vector, and is fed to the decoder which forms the character. However, if the user's previous attempt is 3.5 stars or above then the user's previous drawing is fed as an input to sketch-RNN that encodes it into the latent vector and decodes it into a letter matching the user style. As long as the student performs well, the instruction matches their drawing style but will revert to the "correct" way of writing letters if the student starts underperforming. The personalized instruction can invite learners to interact with the same letter for longer periods, but it does not rise to the level of instructional modeling as realized in Level 5 ITS surveyed in Section 2.1.3.

4 DISCUSSION AND FUTURE WORK

A technical limitation of Djehuty is the small dataset used for training. Every letter has only 120 examples written by 60 different writers. It appears that some letters exhibit a greater diversity than others, and the quality of the dataset might not be suitable for a model supposed to teach handwriting. Likely due to the limited training corpus, the noise introduced in sketch-RNN sometimes led to drawings which were somewhat difficult to identify as the target letter. Moreover, since the user input space is unrestricted, modeling new letters after the user's drawing could lead to unexpected results; therefore, the 3.5 threshold was imposed to avoid using bad input to the decoder. However, training on a larger corpus could improve the compression and we could be more lenient regarding the student's input to the sketch-RNN. Moreover, the educational features and impact of Djehuty should be assessed, e.g. by measuring gain in handwriting fluency and legibility. Further comparison with other digital methods of handwriting teaching will be conducted as performed in [30] and [26], to show the effect of mixed-initiative on handwriting instruction.

5 CONCLUSION

This paper introduced Djehuty, a mixed-initiative handwriting game on the Android platform. The system uses a neural network to produce an infinite variation of letters for children to replicate. Djehuty incorporates gamification features such as simplicity, feedback, real-time response and ITS features such as interactivity, expert model, student model, generativity, and mixed-initiative.

REFERENCES

[1] Felix Albu, Daniela Hagiescu, and Liviu Vladutu. 2015. Intelligent tutor for first grade children's handwriting application. In *Proceedings of the International Technology, Education and Development Conference.*

[2] Kazuaki Ando, Toshinori Yamasaki, and Naoko Yamashita. 2002. Internet-based Japanese language learning system for handwriting Kanji characters beautifully. In *Proceedings of the International Conference on Computers in Education*. IEEE, 867–871.

[3] World Bank. [n.d.]. Literacy rate, adult total (% of people ages 15 and above) | Data. https://data.worldbank.org/indicator/SE.ADT.LITR.ZS?locations=ZG. Accessed 13 Jan 2020.

[4] World Bank. [n.d.]. Population growth (annual %) | Data. https://data.worldbank.org/indicator/SP.POP.GROW?locations=ZG. Accessed 13 Jan 2020.

[5] Hala Bezine and Adel M Alimi. 2013. Development of an Arabic Handwriting Learning Educational System. *International Journal of Software Engineering & Applications* 4, 2 (2013), 33–49.

[6] Elizabeth Bjork and Robert Bjork. 2011. Making things hard on yourself, but in a good way: Creating desirable difficulties to enhance learning. *Psychology and the Real World: Essays Illustrating Fundamental Contributions to Society* (01 2011), 56–64.

[7] Jaime R. Carbonell. 1970. AI in CAI: An Artificial Intelligence Approach to computer Aided Instruction. *IEEE Transactions on Man-Machine Systems* 11 (1970), 190–202.

[8] Junyoung Chung, Çaglar Gülçehre, KyungHyun Cho, and Yoshua Bengio. 2014. Empirical Evaluation of Gated Recurrent Neural Networks on Sequence Modeling. *CoRR* abs/1412.3555 (2014). arXiv:1412.3555 http://arxiv.org/abs/1412.3555

[9] Edward De Bono. 2010. *Lateral thinking: Creativity step by step.* Harper Collins.

[10] Alberto de Diego-Cottinelli and Beatriz Barros. 2010. TRAZO: a tool to acquire handwriting skills using tablet-PC devices. In *Proceedings of the 9th International Conference on Interaction Design and Children - IDC '10*. ACM Press.

[11] Najim Dehak, Reda Dehak, James Glass, Douglas Reynolds, and Patrick Kenny. 2010. Cosine Similarity Scoring without Score Normalization Techniques. *Odyssey* (2010).

[12] Ahmed El-Sawy, Mohamed Loey, and Hazem EL-Bakry. 2016. Arab Kids Tutor (AKT) System For Handwriting Stroke Errors Detection. 4, 3 (2016), 9.

[13] Steve Graham and Naomi Weintraub. 1996. A review of handwriting research: Progress and prospects from 1980 to 1994. *Educational psychology review* 8, 1 (1996), 7–87.

[14] David Ha and Douglas Eck. 2017. A Neural Representation of Sketch Drawings. *arXiv:1704.03477 [cs, stat]* (May 2017). http://arxiv.org/abs/1704.03477 arXiv: 1704.03477.

[15] Deanna Hood, Séverin Lemaignan, and Pierre Dillenbourg. 2015. When Children Teach a Robot to Write: An Autonomous Teachable Humanoid Which Uses Simulated Handwriting. In *Proceedings of the Tenth Annual ACM/IEEE International Conference on Human-Robot Interaction - HRI '15*. ACM Press, 83–90.

[16] Zhihui Hu, Howard Leung, and Yun Xu. 2007. Stroke Correspondence Based on Graph Matching for Detecting Stroke Production Errors in Chinese Character Handwriting. In *Advances in Multimedia Information Processing – PCM 2007*, Horace H.-S. Ip, Oscar C. Au, Howard Leung, Ming-Ting Sun, Wei-Ying Ma, and Shi-Min Hu (Eds.). Vol. 4810. Springer Berlin Heidelberg, 734–743.

[17] Zhi-Hui Hu, Yun Xu, Liu-Sheng Huang, and Howard Leung. 2009. A Chinese Handwriting Education System with Automatic Error Detection. *Journal of Software* 4, 2 (2009).

[18] Nazma Khatun and Jouji Miwa. 2016. An Autonomous Learning System of Bengali Characters Using web-based intelligent Handwriting Recognition. *Journal of Education and Learning* 5, 3 (2016).

[19] Diederik P Kingma and Max Welling. 2013. Auto-Encoding Variational Bayes. arXiv:1312.6114 [stat.ML]

[20] Antonios Liapis, Georgios N. Yannakakis, and Julian Togelius. 2013. Sentient World: Human-Based Procedural Cartography. In *Proceedings of Evolutionary and Biologically Inspired Music, Sound, Art and Design (EvoMusArt)*. Vol. 7834, LNCS. Springer, 180–191.

[21] David Llorens, Federico Prat, Andrés Marzal, Juan Vilar, María José Castro-Bleda, Juan Amengual, Sergio Barrachina Mir, Antonio Castellanos, Salvador Boquera, Jon Gómez, Jorge Gorbe-Moya, Albert Gordo, Vicente Palazón-González, Guillermo Ripollés, Rafael Ramos-Garijo, and Francisco Zamora-Martínez. 2008. The UJIpenchars Database: a Pen-Based Database of Isolated Handwritten Characters.. In *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC'08).*

[22] Faraón Llorens-Largo, Francisco J. Gallego-Durán, Carlos J. Villagrá-Arnedo, Patricia Compañ Rosique, Rosana Satorre-Cuerda, and Rafael Molina-Carmona. 2016. Gamification of the learning process: lessons learned. *IEEE Revista Iberoamericana de Tecnologias del Aprendizaje* 11, 4 (2016), 227–234.

[23] Hieu V. Nguyen and Li Bai. 2011. Cosine Similarity Metric Learning for Face Verification. In *Computer Vision – ACCV 2010 (Lecture Notes in Computer Science)*, Ron Kimmel, Reinhard Klette, and Akihiro Sugimoto (Eds.). Springer, Berlin, Heidelberg, 709–720. https://doi.org/10.1007/978-3-642-19309-5_55

[24] David G. Novick and Stephen Sutton. 1997. What is Mixed-Initiative Interaction?. In *Proceedings of the AAAI Spring Symposium on Computational Models for Mixed Initiative Interaction.*

[25] Kenechi Okeleke and Stefano Suardi. 2019. GSMA Intelligence — Research — The Mobile Economy Sub-Saharan Africa 2019. https://www.gsmaintelligence.com/research/2019/07/the-mobile-economy-sub-saharan-africa-2019/786/. Accessed 13 Jan 2020.

[26] Melissa M. Patchan and Cynthia S. Puranik. 2016. Using tablet computers to teach preschool children to write letters: Exploring the impact of extrinsic and intrinsic feedback. *Computers & Education* 102 (Nov. 2016), 128–137. https://doi.org/10.1016/j.compedu.2016.07.007

[27] Damien Simonnet, Eric Anquetil, and Manuel Bouillon. 2017. Multi-criteria handwriting quality analysis with online fuzzy models. *Pattern Recognition* 69 (2017), 310–324.

[28] UNESCO Institute for Statistics (UIS). 2019. SDG 4 data book: global education indicators 2019. (2019).

[29] UNICEF. 2019. *A world ready to learn prioritizing quality early childhood education.* United Nations Children's Fund, New York.

[30] Sabine Wollscheid, Jørgen Sjaastad, Cathrine Tømte, and Nalini Løver. 2016. The effect of pen and paper or tablet computer on early writing – A pilot study. *Computers & Education* 98 (2016), 70–80. https://doi.org/10.1016/j.compedu.2016.03.008

[31] Beverly Park Woolf. 2009. *Building intelligent interactive tutors: student-centered strategies for revolutionizing e-learning.* Morgan Kaufmann Publishers/Elsevier.

[32] Georgios N. Yannakakis, Antonios Liapis, and Constantine Alexopoulos. 2014. Mixed-initiative co-creativity.. In *FDG*.