מערכות דינמיות 2 106970 – מגישה: זרוהי פפיאן 316648799

רשת נוירונים וכאוס:

בנייה, זיהוי וחיזוי של כאוס עם רשת נוירונים מלאכותית רב שכבתית

כיום ניתן לראות כי המושג "כאוס", או כאוטי, נמצאים בהרבה מתחומי החיים. בתחום מדעי המוח מצאו כי קיימת דינמיקה כאוטית, ולא רק בנוירונים - אלא גם ברשתות עצביות במוח.

המושג "רשתות נוירוניות כאוטיות", אותו נגדיר בהמשך, נפוץ גם בתחום המחשבים. הדינאמיקה המורכבת מאפשרת שימושים, כגון: זיכרונות אסוציאטיביים, כלי אבטחה דיגיטליים(פונקציות *hash*), תוכניות הצפנה וכו'.

הסיבה לשימוש - שליטה **בדינמיקה כאוטית** לשימור דפוסים, עם היתרון העיקרי של אופציית אחסון גדולה.

מוטיבציה נוספת לשימוש היא - חוסר היציבות של המערכת וההתנהגות האקראית והלא צפויה.

אבל הבעיה היא - כי לרוב, הטיעון כי הרשת היא אכן כאוטית, הוא כלל ללא הוכחה מתמטית.

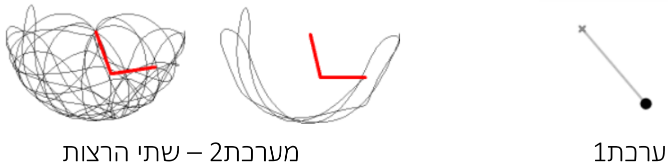
בעזרת ההגדרה של $Devaney$ עבור כאוס – המאמר הראה כי אפשר להשתמש בכלים מתמטיים כדי **להעריך את ההתנהגות של המערכת הדינמית**. בנוסף, המאמר הצליח למצוא שיוויון מסוים בין איטרציות כאוטיות לפי $Devaney$ וקבוצה מסוימת של רשתות נוירוניות MLP.

נחקרה גם הבעיה ההפוכה – חקירת היכולת של MLP ללמוד *משפחה מסוימת של מערכות דינמיות כאוטיות.*

רקע – נזכר במספר הגדרות:

1. **מערכת דינמית** - זוג (χ, f) כאשר χ מרחב מטרי קומפקטי עם המטריקה d והעתקה רציפה $f : \chi \to \chi$
2. **איטרציה של מערכת דינמית** – הצעד המוכל בעדכון המצב הגלובלי x^t בזמן t ביחס לפונקציה f כך ש $x^{t+1} = f(x^t)$. (עם פעולת הרכבה)
3. **מערכת דינמית כאוטית** לפי הגדרת *הכאוס הדטרמיניסטי* של $Devanry$ –

ראשית, נבין את המושג – *כאוס דטרמיניסטי*[1]. מנקודת מבט ראשונה, זה נשמע יותר כמו אוקסימורון – כי המילה כאוס מתקשרת יותר עם המילה רנדומליות. אבל אפילו משהו רנדומלי יכול לעקוב אחרי דפוס פשוט, כך שהאקראיות לבד היא לא תנאי מספיק. למעשה, זה אפילו לא הכרחי. כדי להבין יותר טוב נעזר בדוגמה – ניקח שתי מערכות עם *מטוטלת*[2] ולאחר ההרצה ראשונה נשנה במעט את המיקום הראשוני של כל אחת מהן:



מערכת2 – שתי הרצות   מערכת1

<u>מערכת1</u>: מטוטלת אחת – קל לראות כי שינוי קטן במצב הראשוני יוביל לשינוי קל באופן שבו המערכת תתפתח. אולי נקבל תנועה קצת שונה, אבל המסלולים הכוללים ייראו כמעט אותו דבר

<u>מערכת2</u>: מטוטלת כפולה[3] – למרות שאפשר יהיה להבחין בדפוסים מסוימים, נקבל מסלול שונה לחלוטין. במילים אחרות, מערכת2 היא **רגישה** מאוד לתנאים ההתחלתיים.

[1] **דטרמיניסטי** - מהמילה $determine$ = לקבוע

[2] **מטוטלת** = מוט אשר בקצה האחד מחובר משקל ובקצה השני מחובר היטב למשהו מקרקעין

[3] **מטוטלת כפולה** = 2 מטוטלות מחוברות כך שהן יכולות להתקפל

זהו המאפיין החשוב ביותר של מערכת כאוטית: מושג המייצג בד"כ את **אפקט-הפרפר**:

" שינויים קטנים בתנאים הראשוניים עשויים להוביל להתנהגויות שונות "

[כלומר, גם שני מחשבים המחשבים את אותה משימה בשיטות דומות (אך לא שוות) עם דיוק דומה יגיעו לפתרונות <u>שונים</u>, משום ש"התנאים ההתחלתיים "יהיו שונים במקצת בכל שלב של הסימולציה]

דוגמה נוספת היא *מודלים למזג האוויר*. המודלים מביאים תחזית המתאימה רק לשבוע, כי הם מאותחלים עם תנאים המשקפים רק ידע חלקי של מזג האוויר הנוכחי. גם אם ההבדלים בהתחלה קטנים, הם מתרבים במרוצת הזמן, ובכך מקשים על חיזוי לטווח ארוך.

למעשה, במערכת כאוטית, קירוב של המצב הנוכחי הוא **אינדיקטור חסר תועלת למדי** לניבוי עתידיות.

ההגדרה הפורמלית עבור **- הכאוס הדטרמיניסטי** לפי $Devaney$, מורכבת משלושה תנאים:

תהי $f : \chi \to \chi$. היא כאוטית על χ אם:

1. **טרנזיטיביות טופולוגית**[4] [לא ניתנת לצמצום]
2. קבוצת **הנקודות-המחזוריות**[5] **– צפופה**[6] ב χ [ז"א ההעתקה רגולרית]
3. f בעלת **תלות נקודתית רגישה בתנאים ההתחלתיים**[7]

נשים לב כי מערכת דינמית העושה איטרציות על f תקרא **כאוטית** לפי $Devaney$ על (χ,τ) אם –

f מקיימת את שלושת התכונות, כאשר:

1. מרמז כי מערכת דינמית לא יכולה להיות מחולקת ("שבורה") לתת מערכות פשוטות יותר
2. הכאוס צריך רגולציה כדי "לנגד" את ההשפעות של הטרנזיטיביות, וזהו אלמנט הקביעות
3. מערכת לא צפויה - לכן, שתי נקודות הקרובות אחת לשנייה יכולות להתנהג בצורה שונה לגמרי

מה שגורר כי *האיטרציות* נקראות **כאוטיות** (על פי $Devaney$) כאשר *המערכת הדינמית* היא כאוטית.

נסתכל כעת על **תחום האיטרציות** - עד כמה שידוע לנו, אין כל *תוצאת מבחן* שההתנהגות הכאוטית של מערכת דינמית שהוכחה, באופן תיאורטי, על $\mathbb{R}$ נשארת תקפה גם **בנקודות הצפה** [$floating\ point$], שהיא תחום היישום. לכן, כדי למנוע אובדן של כאוס, המאמר מציג *תחלופה* - שהיא לעשות איטרציה של **העתקות בוליאניות**: התוצאות שהושגו תיאורטית בתחום זה נשמרים ביישומים.

[במחשב יש מספר סופי, וכדי להימנע מעניין ספרות הדיוק, מבעיה נומרית - משתמשים ב0,1]

מספר סימונים חדשים:

- $[[\mathbf{1};\boldsymbol{n}]]$ - המרחק בין המספרים $\{1,1+1,\dots,n\}$
  במערכת כל רכיב $i \in [[1;n]]$ לוקח את הערך שלו x_i ב $\mathbb{B}=\{0,1\}$
- **תצורת המערכת בזמן דיסקרטי $\boldsymbol{t}$** - היא הווקטור $x^t=(x_1^t,\dots,x_n^t)\in\mathbb{B}^n$
- **דינמיקת המערכת** - מוגדרת לפי הפונקציה $f:\mathbb{B}^n\to\mathbb{B}^n$ כך ש $f(x)=\big(f_1(x),\dots,f_n(x)\big)$
- **תצורת $\boldsymbol{x}$**: $\boldsymbol{N(i,x)=(x_1,\dots,\overline{x_i},\dots,x_n)}$ ($\overline{x_i}$ ההפך של x_i). [אינטואיטיבית, $N(i,x)$,x הם שכנים]

[4] טרנזטיבית טופולוגית - $\forall U,V\in\chi\ (open):\ \exists k>0\ s.t.\ f^{(k)}(U)\cap V=\emptyset$

[5] x נקודה **מחזורית** עבור f מסדר $n\in\mathbb{N}^*$ אם $f^n(x)=x$

[6] f **צפופה** בX אם לכל $x\in X$ אפשר למצוא לפחות נקודה מחזורית אחת באחת הסביבות

[7] f בעלת **תלות נקודתית רגישה בתנאים ההתחלתיים**, ז"א [δ קבוע רגישות]
$\exists\delta>0\ s.t.\ \forall x\in X\ \&\ any$ neighbourhood $V_x,\ \exists y\in V_x\ n>0\ s.t\ \ d\big(f^n(x),f^n(y)\big)>\delta$

- $\Gamma(f)$ = **גרף אוריינטבילי של איטראציות** (בדידות של f) כאשר בגרף כזה:
  - קודקודים הם תצורות של $\mathbb{B}^n$ - יש קשת-i בין $x, N(i,x)$ אם ורק אם $f_i(x) = N(i,x)$

- **אסטרטגיה** $S = (S^t)^{t\in\mathbb{N}}$ - רצף המגדיר איזה רכיב יש לעדכן בזמן t, כאשר S^t מסמן את התנאי הt

- **איטרציה אסינכרונית** - סכמת איטרציות המשנה רכיב **אחד** בלבד בכל איטרציה. באופן מדויק יותר: עבור $i \in [1, n]$ כלשהו

$$(1) \quad f(x) = \begin{cases} x^0 \in \mathbb{B}^n \\ x_i^{t+1} = \begin{cases} f_i(x^t), & S^t = i \\ x_i^t \ , & else \end{cases} \end{cases}$$

- הפונקציה $F_f: [[1;n]] \times \mathbb{B}^n \to \mathbb{B}^n$ המוגדרת לכל $f: \mathbb{B}^n \to \mathbb{B}^n$ כך ש:

$$(2) \quad F_f(s,x)_j = \begin{cases} f_j(x), & j = s \\ x_j \ , & else \end{cases} \quad , \quad (3) \quad \begin{cases} x^0 \in \mathbb{B}^n \\ x^{t+1} = F_f(S^t, x^t) \end{cases}$$

- איטרציות המוגדרות ב(3) יכולות להיות מוסברות ע"י המערכת הבאה (עם סימון חדש - G_f):

$$(4) \quad \begin{cases} X^0 = \left((S^t)^{t\in\mathbb{N}}, x^0\right) \in [[1;n]]^{\mathbb{N}} \times \mathbb{B}^n \\ X^{k+1} = G_f\left(X^k\right) \end{cases} \qquad , G_f\left(\left((S^t)^{t\in\mathbb{N}}, x\right)\right) = \left(\sigma\left((S^t)^{t\in\mathbb{N}}\right), F_f(S^0, x)\right)$$

כאשר σ היא פונקציית ה$shift$ הידועה המוציאה את האיבר הראשון של האסטרטגיה (S^0)

הגדרה זו מאפשרת לקשר בין *איטרציה אסינכרונית* עם *האיטרציה הקלאסית של המערכת הדינמית* (G_f).

נשים לב כי G_f כאוטית לפי $Devaney$.

כעת, נרצה ללמוד **תכונות טופולוגיות** של האיטרציות, ולשם כך נגדיר - מרחק d בין שתי נקודות:

$$(5) \quad d\left((S,x);(\check{S},\tilde{x})\right) = d_e(x,\check{x}) + d_s(S,\check{S})$$

$$(6)^{[8]} \quad d_e(x,\check{x}) = \sum_{j=1}^{n} \Delta(x_j, \check{x}_j) \in [[0;n]] \qquad (7) \quad d_s(S,\check{S}) = \frac{9}{2n}\sum_{t=0}^{\infty}\frac{|S^t - \check{S}^t|}{10^{t+1}} \in [0;1]$$

מרחק זה משקף את המידע הבא – עבור שתי המערכות:

- הבעלות רכיבים ואסטרטגיות דומות, שיש להן אותם תנאי ההתחלה, חייבות להניב מרחק קטן
- ככל שהרכיבים של המערכות שונים, כך גדל המרחק ביניהם

בשלב הבא, נרצה להבין מה היא רשת נוירונים כאוטית :

כדי להבין מהי רשת נוירונים נגדיר קודם מהו *נוירון מלאכותי* בשם **פרספטרון** [$perceptron$] -

הפרספטרון מקבל קלט בינארי (ברבים) ומייצר קלט בינארי יחיד.

רשת נוירונים מלאכותית [ANN] – זהו מודל מתמטי חישובי שפותח בהשראת התהליכים הקוגניטיביים במוח במסגרת למידת מכונה. הרשת מכילה בדרך כלל מספר רב של יחידות מידע (קלט ופלט) מקושרות. כאשר צורת הקישור בין היחידות, המכילה מידע על חוזק הקשר, מדמה את אופן חיבור הנוירונים במוח.

פרספטרון רב שכבתי [MLP] – מכיל יותר משתי שכבות אימון (קלט, פלט) וכל שכבה נוספת, בין לבין, נחשבת ל"חבויה". ברשת זו כל החצים הולכים לכיוון של הפלט בלבד.

[8] Δ - $\begin{cases} \Delta(x,y) = 0, & x = y \\ \Delta(x,y) = 1, & else \end{cases}$

רשת נוירונים, הנבנית על ידי שילוב של **פונקציות ההעברה ותנאים התחלתיים** – ששניהם כאוטיים, נחשבת גם לכאוטית. רשתות נוירוניות כאוטיות נבנו בגישות שונות;

- בהקשר של **זיכרון אסוציאטיבי**, נוירונים כאוטיים, כמו נוירונים של **מצב דינמי לא לינארי**, לעתים קרובות מהווים את הצמתים של הרשת. לנוירונים אלה יש התנהגות כאוטית מובנית ($inherent$), אשר מוערכת בדרך כלל באמצעות חישוב של **המעריך Lyapunov**

- גישה חלופית, אותה נראה, היא לשקול מבנה ידוע של רשת נוירונים MLP . דבר המתאים למדל יחסים **לא לינאריים** בין נתונים, בשל *יכולת הקירוב האוניברסלית*[9] שלהן. לכן, סוג זה של רשתות ניתן לאמן למדל תופעה פיזית, הידוע בתור כאוטית, בשם "מעגל של $Chua$"

לבניית רשת נוירונים כאוטיות לפי $Devaney$ נעזר במשפט וביחס הבאים:

(משפט 1) תהי $f: \mathbb{B}^n \to \mathbb{B}^n$. איטרציות של G_f הן כאוטיות לפי $Devaney \Leftrightarrow \Gamma(f)$ **קשיר בהחלט**[10]

היחס בין $\Gamma(f), G_f$ – קיים מסלול בין x, x' ב $\Gamma(f) \Leftrightarrow$ קיימת אסטרטגיה S כך שהאיטרציות של G_f מהנקודה הראשונית (s, x) מגיעות לקונפיגורציה x'. $(Guyeux)$

כעת, נגדיר 2 פונקציות $f_0, f_1 : \mathbb{B}^n \to \mathbb{B}^n$ בצורה הבאה: (נראה אותן גם בהמשך בדוגמה)

$$f_0(x_1, \dots, x_n) = (\overline{x_1}, \dots, \overline{x_n})$$
$$f_1(x_1, \dots, x_n) = (\overline{x_1}, x_1, x_2, \dots, x_{n-1})$$

קל לראות כי $\Gamma(f_0), \Gamma(f_1)$ קשירים בהחלט, אז האיטרציות של G_{f_0}, G_{f_1} הן כאוטיות לפי $Devaney$.

בניית רשת MLP הממדלת $F_{f_0} : \chi \to \mathbb{B}^n$[11] קשורה **לשליליות וקטורית** (ז"א $\overline{x_i}$).

כעת, אפשר לקשר בין השכבות של הקלט והפלט כדי להדגים את התלות בין שתי איטרציות חוזרות.

אנו מקבלים **רשת נוירונית גלובלית חוזרת** המתנהגת באופן הבא:

- אתחול עם ווקטור הקלט $(S^0, x^0) \in \chi$ המחשב את ווקטור הפלט $x^1 = F_{f_0}(S^0, x^0)$ הנשלח בחזרה לשכבת הקלט דרך **הקישורים בפידבק**[12]

- בהפעלת האיטרציה הt^{th}, מצב המערכת $x^t \in \mathbb{B}^n$ המתקבל משכבת הפלט ותנאי ההתחלתה של הרצף $(S^t)^{t \in \mathbb{N}}$, משמשים לחישוב ווקטור הפלט החדש, המייצג את המצב החדש של המערכת:
$$(8) \quad x^{t+1} = F_{f_0}(S^0, x^t) \in \mathbb{B}^n$$

ההתנהגות של רשת הנוירונים היא כזו שכאשר המצב הראשוני הוא $x^0 \in \mathbb{B}^n$ והרצף $(S^t)^{t \in \mathbb{N}}$ ניתן כקלט-חיצוני, אז רצף ווקטורי הפלט $(x^t)^{t \in \mathbb{N}^*}$ הוא בדיוק הזה שיוצר בהרצות הכאוטיות המיוצגות ב (4).

משמעות הדבר היא כי <u>מבחינה מתמטית</u>, במידה ווקטורי הקלט דומים, אז שניהם יוצרים את אותם הפלטים $(x^t)^{t \in \mathbb{N}^*}$, ולכן הם רפורמולציות זהות של האיטרציות של G_{f_0} בχ.

לבסוף, כיוון שרשת הנוירונים המוצעת **בנויה על מנת למדל את ההתנהגות שלG_{f_0}**, שהאיטרציות שלה כאוטיות לפי $Devaney$, ניתן להסיק שגם הרשת היא כאוטית.

[9] לפי **משפט הקירוב האוניברסלי** לכל פונקציה אשר ממפה רצף של מספרים ממשיים לרצף אחר של מספרים ממשיים ניתן למצוא קירוב טוב ככל שנרצה על ידי פרספטרון רב-שכבתי שבו רק שכבה חבויה אחת.

[10] לבדוק אם הגרף קשיר בהחלט לא קשה - אלגוריתם $Tarjan$ למרחק

[11] סימון: $\chi = [\![1; n]\!] \times \mathbb{B}^n$

[12] רשת נוירונים המציגה לולאה נקראת חוזרת או **פידבק** – כאשר הכוונה פה לפידבק אחורה

למרות השימוש בf_0, נשים לב כי זה לא מוגבל. ניתן להרחיב לכל פונקציה f כך ש G_f היא העתקה כאוטית על ידי אימון הרשת למודל $F_f : \chi \rightarrow \mathbb{B}^n$. וממשפט 1, אנו יכולים למצוא פונקציות חלופיות f עבור f_0 דרך בדיקה פשוטה של גרף האיטרציות $\Gamma(f)$. (לבנות עוד רשת נוירונית כאוטית באמצעות f_1 במקום f_0)

כעת, לאחר שהצלחנו *לבנות*, נרצה לדעת גם איך ניתן *<u>לבדוק</u>* האם הרשת הנתונה היא כאוטית או לא. כמו שצוין לפני, בהרבה מהמאמרים הרשת הנוירוניות נחשבת לכאוטית ללא כל הוכחה מתמטית משכנעת.
המאמר הציג גישה להתגבר על זה עבור קטגוריה מסוימת של MLP אותה הגדרנו לפני כן בבניית הרשת. [13]:

אפשר להוכיח כי ההתנהגות הטופולוגית של רשתות נוירוניות אלו היא *כאוטית* בתהליך הבא;

- נסמן $F : \chi \rightarrow \mathbb{B}^n$ [11] פונקציה הממפה את הערך $(s,x) \in \chi$ לערך $y \in \mathbb{B}^n$
- נסמן $f : \mathbb{B}^n \rightarrow \mathbb{B}^n$ כך ש: $f(x) = \big(F(1,x)), \dots, F(n,x)\big)$ (9)
  לכן, עבור j כלשהו, $j \in [1,n]$, יש לנו $f_j(x) = F(j,x)$

אם רשת נוירונים חוזרת זו מתחילה עם (S, x^0), זה מייצר בדיוק את אותו פלט ווקטורים שהאיטרציות הכאוטית של F_f עם המצב הראשוני - $(S,x^0) \in \big[[1;n]\big]^{\mathbb{N}} \times \mathbb{B}^n$. במילים אחרות, ווקטורי הפלט של MLP תואמים את רצף הקונפיגורציות הנתון ע"י (3), ומסומנים בתור $\mathbf{CI-MLP}(\boldsymbol{f})$ [14].
בדיקה האם $\mathrm{CI-MLP}(f)$ מתנהגת באופן כאוטי לפי *Devaney* היא פשוטה: יש לבדוק האם גרף האיטרציות האסוציאטיבי ל $\Gamma(f)$ קשיר בהחלט או לא. כתוצאה מקרית, מתקבל שוויון בין איטרציות כאוטיות ו $\mathrm{CI-MLP}(f)$. לכן, אפשר יהיה להשתמש בכלים מתמטיים להבנת ההתנהגות.

לסיום, בהקשר של מדעי המחשב, לתחומי נושא שונים יש עניין בכאוס - כמו טכניקת **סטגנוגרפיה**[15].
בין האסטרטגיות לגלאים וזיהוי המסר הנסתר, קיימים: מכונות תמיכה ווקטוריות, *רשתות נוירוניות* ורשתות מרקוב. כאשר רוב הגלאים מביאים תוצאות די טובות. עם זאת, נדגיש כי אף אחת מצורות ההסתרה שהוצעו לא מקיימת את הגדרת הכאוס לפי *Devaney*. לכן באה השאלה - האם הגלאים ימשיכו להניב תוצאות טובות כאשר יתמודדו עם סכמה כאוטית באמת. באופן כללי יותר, נותרה הבעיה הפתוחה של קביעה האם הבינה המלאכותית מתאימה לחזות התנהגויות טופולוגיות כאוטיות.

לפי מאמרו של כריסטוף גויקס – הסתרת נתונים היא מאובטחת כאשר התנהגותה אינה ניתנת לחיזוי. כאשר מבחינה מתמטית, זהו כאוס לפי *Devaney* עם שלושת התכונות שראינו, עליו הוא גם פירט במאמר.
אלו תכונות שימושיות בהקשר של הגנה והתקפה כי; מועילות לאימות, מסתירות טוב יותר נתונים, משפרות את האבטחה, כי אי אפשר לקבל הבנה טובה יותר של המצב מהסיבה שאי אפשר לצמצם את המורכבות על ידי לימוד רק של צמצום מופרז של התוכן. מוצע המושג החדש של **בטחון-כאוטי**.
הרעיון הוא שתוכנית הסתרת נתונים מסוימת תהיה בטוחה יותר מאחרת אם היא תציג מספר גדול יותר של תכונות כאוטיות, ואם הערכים הכמותיים שלה יהיו טובים יותר.

[13] למרות הגבלה, צוין כי הכותבים סבורים כי ניתן להרחיב גישה זו למגוון רחב של רשתות נוירוניות
[14] $\mathbf{CI-MLP}(\boldsymbol{f}) = $ "*Chaotic Iterations based MultiLayer Perceptron*" - רשתות נוירוניות חוזרות ונשנות
[15] **סטגנוגרפיה** היא אמנות הסתרת המסר. הידע בתוך קובץ המוביל להיכנס כך לקשר אנשי צוות לא מורשים אינו מסוגל לזהות נוכחות של ידע בתוך קובץ המוביל המסר המיועד להסתרה מוטמע בשיטות שונות

נראה כעת **דוגמה להמחשה** של מה שראינו –

נזכר בשתי הפונקציות שהגדרנו: $f_0, f_1 : \mathbb{B}^n \to \mathbb{B}^n$:

$$f_0(x_1, \dots, x_n) = (\overline{x_1}, \dots, \overline{x_n})$$
$$f_1(x_1, \dots, x_n) = (\overline{x_1}, x_1, x_2, \dots, x_{n-1})$$

עכשיו אפשר לבנות רשתות כאוטיות על-ידי למידה של F_{f_0} או של F_{f_1}.

דוגמאות לכאלו:

$$f_0(x_1, x_2, x_3, x_4) = (\overline{x_1}, \overline{x_2}, \overline{x_3}, \overline{x_4})$$
$$f_1(x_1, x_2, x_3) = (\overline{x_1}, x_1, x_2)$$
$$f_2(x_1, x_2, x_3) = (x_1, \overline{x_2}, x_2)$$

אבל, נשים לב כי עבור ההגדרה הבאה של g – הבאים לא כאוטיים:

$$g_0(x_1, x_2, x_3) = (x_1, x_2, x_3)$$
$$g_1(x_1, x_2, x_3) = (\overline{x_1}, x_2, x_3)$$

בנוסף, לפי מאמר נוסף שלהם שפורסם בצרפת[16], נעשה ניסוי – לכל רשת נוירונים חישבו את הערכים:

- *ממוצע מספר האיטרציות האסינכרוניות* - שצריך כדי להכשיר
- *שיעור הצלחה* - המשקף אימון מוצלח בתוך פחות מ 1000 איטרציות

התוצאות שהתקבלו הדגישו מספר נקודות:

ראשית, ארכיטקטורות MLP שונות יכולות ללמוד פונקציית איטרציה זהה. בפרט, הניסוי בדק את ההבדל בין שתי רשתות נוירוניות - רשת עם שכבה מוסתרת אחת בלבד ורשת עם שתי שכבות מוסתרות. התוצאות הראו כי ככל הנראה הרשת עם שתי השכבות המוסתרות הייתה מורכבת מדי ללמידה עבור הפונקציות.

שנית, זה היה נראה כי לאמן רשתות המתנהגות באופן כאוטי, זה קשה יותר, כי הם צריכים, בממוצע, יותר אימונים כדי לאמן כראוי (וככל הנראה יותר דוגמאות כדי לא להגיע למצב של **התאמת יתר**[17]).

עם זאת, עדיין בוחנים את הנקודה הזאת ואין מסקנות סופיות לגבי זה.

שאלה פתוחה נוספת היא - האם יש קשר בין הקושי באימון לבין ההפרעה[18] המושרה על ידי פונקציה איטרציה כאוטית.

שימוש במאמרים הבאים:

1. "Neural Networks and Chaos: Construction, Evaluation of Chaotic Networks, and Prediction of Chaos with Multilayer Feedforward Networks"
   By Jacques M. Bahi, Jean-Fran_cois Couchot, Christophe Guyeux, and Michel Salomon
   August 23, 2016

2. "A chaos-based approach for information hiding security"
   By Jacques M. Bahi, Senior Member IEEE, Christophe Guyeux
   May 2010

[16] מספר 11 ברפרנס של המאמר

[17] **התאמת יתר** – זוהי בעיה יסודית בלמידה שבה המודל מותאם יתר על המידה לאוסף מסוים של נתונים ועל בן מצליח פחות בבצוע תחזיות

[18] **בהפרעה** הכוונה היא להערכה של קבועי הרגישות, התפשטות, וכו '