מעבדה 2 בבינה מלאכותית

נושא:

אלגוריתמים גנטיים, ממטיים ואבולוציה

­­

מגישים:

1. אילן גודיק, 316315332
2. יובל אלפסי, 318401015

מנחה:

מר שי בושנסקי



תאריך הגשה:

8 לאפריל 2016

**תוכן עיניינים**

חלק א: שכלול המנוע הגנטי

שאלה 1 – עמוד ­­\_ – דיווח זמן ריצה

שאלה 2 – עמוד \_ – שיטות בחירה ושיטות שרידות

שאלה 3 – עמוד \_ – פונקציית מרחק בין גנים

שאלה 4 – עמוד \_ – אבחון מינימום מקומי

שאלה 5 – עמוד \_ – היחלצות ממינימום מקומי

שאלה 6 – עמוד \_ – בחינת ההשפעה על האלגוריתם הגנטי

שאלה 7 – עמוד \_ – חזית פרטו אופטימל

חלק ב: אפקט בולדווין

חלק ג': ניתוחים וסטטיסטיקה

**המנוע הגנטי**

**ארכיטקטורה**

תחילה נתאר את מבנה המנוע הגנטי החדש, תהליך האבולוציה והרכיבים הכלולים בכך.

Generation:

- Fitness Mapping Functions

- Breeding: Population => Stream[Child]

- Choose Parents: Population => Parent

- Crossover: Parent, Parent => Child

- Mutation: Child => Mutate

- Survivor Selection: Population, Children => new Population

Genetic Engine:

- Local Optima Detection

- Normal Generation

- Local Optima Generation

Genetic Alg:

- Genetic Problem & Representation

- Genetic Engine

**ייצוג הבעייה:** **Genetic Problem & Representation**

על מנת לממש תמיכה בבעיה חדשה ופתרונה ע"י המנוע הגנטי, דרוש רק מימוש של הממשק הבא וייצוג פרמטרי שלה (הסבר בהמשך)

**public interface** Genetic<A> {  
 **double** fitness(A gene);  
 A mate(A x, A y);

*// Allowed to be in place (done on fresh gene generated by mate)* A mutate(A a);

Metric<A> metric();  
 A randomElement(Random rand);  
 String show(A gene);  
}

וזה כל מה שצריך לממש על מנת להפעיל את המנוע הגנטי על בעייה חדשה.

באופן אוטומטי, ניתן להריץ על הבעיה את כל הטכניקות של האלגוריתם הגנטי, לבחור אילו שיטות גנטיות להפעיל, כגון Local Optima Detection, כל פונקציות הFitness Mapping, להחליף את שיטת הבחירה, המוטציה וההישרדות, ואף ניתן להפעיל את האלגוריתם הMeta Genetic, על המנוע ועל הפרמטרים הספציפיים לבעיה.

**שאלה 1:**

הוספנו הדפסה של זמן הריצה של האלגוריתם:

כמו כן, ב-JVM אין תמיכה בגישה לכמות מעגלי השעון של המעבד, לפיכך לא יכולנו להוסיף תמיכה בכך.

דוגמאת הרצה: (חיפוש מחרוזת באלגוריתם גנטי)



**שאלה 2:**

RWS: אלגוריתם הRWS דוגם את הגנים בהסתברות – יחס ישיר לfitness.

מימשנו את אלגוריתם הבחירה RWS בעזרת Stochastic Acceptance שרץ בזמן :

# Roulette-wheel selection via stochastic acceptance [[Lipowski](http://arxiv.org/find/cs/1/au:+Lipowski_A/0/1/0/all/0/1), [Lipowska](http://arxiv.org/find/cs/1/au:+Lipowska_D/0/1/0/all/0/1), 2011]

***http://arxiv.org/abs/1109.3627***

קוד:

**public** <A> A chooseSingleParent(Population<A> population, Random rand) {  
 **int** popSize = population.**population**.**length**;  
 **double** maxFitness = *maxFitness*(population);  
 **int** index;  
 **do** {  
 index = rand.nextInt(popSize); *// 1/N probability to choose anyone* } **while** (rand.nextDouble() >=

(1 - population.**population**[index].**fitness**) / maxFitness);  
 **return** population.**population**[index].**gene**;  
}

\* נצפו שיפורים טובים במיוחד בביצועים בעזרת RWS בבעיית מציאת המינימום לפונקציה, בשתי הפונקציות.

השוואה של שיטה תחת פרמטים אופטימיים (ע"פ האלגוריתם הMeta-Genetic), עם בדיקת יציבות ע"י ממוצע של 20,000 הרצות של האלגוריתם הגנטי: (זמן התכנסות לפתרון האופטימלי)

**Top Selection:** 0.55 ms

**RWS:** 0.35 ms

**SUS:** 0.45 ms

**Tournament:** 0.45 ms

Ranking: בחירה כמו בRWS, אך אנחנו מתחשבים אך ורק בטיב היחסי של הגן במערך ממוין:

כך אנחנו נותנים הזדמנות גבוהה יותר לגנים לווא דווקא טובים להיבחר,

ובנוסף, אם פונקציית הfitness לא מאוזנת וההבדל בין fitness של גנים גדול מאוד, אז במקום שתהיה העדפה מאוד חזקה לקצת הגנים הטובים, כעת לכולם יש הסתברות יותר הוגנת להיבחרות, ובכך אנחנו מגדילים את הDiversity.

קוד:

**public** <A> A chooseSingleParent(Population<A> population, Random rand) {  
 **int** popSize = population.**population**.**length**;  
 **double** maxFitness = popSize - 1;  
 **int** index;  
 **do** {  
 index = rand.nextInt(popSize); *// 1/N probability to choose anyone* } **while** (rand.nextDouble() >= index / maxFitness);  
 **return** population.**population**[index].**gene**;  
}

SUS:

SUS מבצע בחירה של גנים בהתאם לfitness שלהם, ומספר הבחירות פרופורציוני בדיוק לfitness כך שאם יש לנו fitnessים של 0.5,0.25,0.25 אז ייבחרו בדיוק 50% אבות מהגן הראשון, בדיוק 25% מהגן השני ו25% מהגן השלישי.

בעזרת אלגוריתם בחירה זה, אנו מבטיחים הופעה גם של גנים פחות טובים, כל עוד האוכלוסייה שאנו בוחרים גדולה מספיק, ובנוסף ייתכן שגן נדיר אף יותר ייבחר אם אנחנו פגענו בסיבוב ההתחלתי הראשון במיקום, כך שבהזזות של 1/size נגיע אליו.

קוד:

**public** <A> Supplier<A> chooseParents(Population<A> population, **int** size, Random rand) {  
 *// Calculate the sum of all fitness values.* **double** fitnessSum = 0;  
 **for** (Gene<A> candidate : population.**population**) {  
 fitnessSum += (1 - candidate.**fitness**);  
 }  
 **double** startPos = rand.nextDouble() \* fitnessSum;  
 RoulettePosition initialPosition = *consume*(startPos, **new** RoulettePosition(0, 1 - population.**population**[0].**fitness**), population);  
 **double** stepSize = fitnessSum / size;  
 Supplier<A> supplier = **new** Supplier<A>() {  
 RoulettePosition **pos** = initialPosition;  
 @Override  
 **public** A get() {  
 **int** i = **pos**.**index**;  
 **pos** = *consume*(stepSize, **pos**, population);  
 **return** population.**population**[i].**gene**;  
 }  
 };  
 List<A> parents = Stream.*generate*(supplier).limit(size).collect(Collectors.*toList*());  
 Collections.*shuffle*(parents);  
 Iterator<A> iterator = parents.iterator();  
 **return** iterator::next;  
}

**private static** <A> RoulettePosition consume(**double** amount, RoulettePosition pos, Population<A> pop) {  
 **int** i = pos.**index**;  
 **double** remaining = pos.**remaining**;  
 **while** (remaining <= amount) {  
 **int** nextIndex = (i + 1) % pop.**population**.**length**;  
 amount = amount - remaining;  
 i = nextIndex;  
 remaining = 1 - pop.**population**[nextIndex].**fitness**;  
 }  
 **return new** RoulettePosition(i, remaining - amount);  
}

Scaling:

מימשנו 3 שיטות Scaling:

1. Windowing – הפחתת הfitness הגרוע ביותר, אך לא עד ל1 (הכי גרוע) כי אז האפשרות הזו לעולם לא תבחר בRWS, וזאת בכדי ליצור בחירה יותר אחידה והוגנת, ובכך אנו מגדילים את הDiversity.

קוד:

**public** <A> **double** mapFitness(Metric<A> metric, Population<A> population, Gene<A> gene) {  
 *// return 1 - ((1 - gene.fitness) - (1 - population.worstMaxFitness()));* **if** (gene.**fitness** == 0)  
 **return** 0;  
 **else** {  
 **double** fitness = 1 + gene.**fitness** - population.worstMaxFitness() - epsilon;  
 **return** Math.*max*(fitness, ***epsilon***);  
 }  
}

1. Exponential Scaling – שורש הfitness, כדי להפחית את השפעתם הקיצונית של הגנים בעלי fitness טוב מאוד.

קוד:

**public** <A> **double** mapFitness(Metric<A> metric, Population<A> population, Gene<A> gene) {  
 **return** 1 - Math.*sqrt*(1 - gene.**fitness**);  
}

1. Sigma Scaling – התחשבות במרחק מהתוחלת ביחס לסטיית התקן, כדי לדרוש Diversity – שיהיו לנו גנים שונים ככל האפשר

קוד:

**public** <A> **double** mapFitness(Metric<A> metric, Population<A> population, Gene<A> gene) {  
 **if**(gene.**fitness** == 0.0) **return** 0;  
 **if**(population.fitnessStdDev() == 0) **return** 0.99;  
 *// 1/g such that it's a minimization problem.* **return** Math.*min*(0.99, Math.*max*(***epsilon***,

1 / (Math.*abs*(gene.**fitness** - population.fitnessAvg()) / 2 \*

population.fitnessStdDev())));  
}

Aging Model:

לכל גן נשמור את הגיל שלו, כך שהוא נולד בגיל 0, ובכל דור גילו עולה ב1.

Aging משנה את הfitness כך שגנים צעירים מקבלים penalty, ככל שהם מתבגרים עד לגיל הבגרות הfitness שלהם משתפר, ומאז ככל שהם מזדקנים, יחס הfitness קטן חזרה לגיל הצעירים.

הFitness Mapping של Aging מתבסס על שתי פרמטרים:

**public final int matureAt**;  
*// Influence: multiplied by a number in the range [1-agingInfluence, 1]***public final double agingInfluence**;

קוד:

**public** <A> **double** mapFitness(Metric<A> metric, Population<A> population, Gene<A> gene) {  
 **int** age = gene.**age**;  
 **if**(**matureAt** == 0) **return** age;  
 *// Scale x such that matureAt goes to 0.5.* **else  
 return** Math.*max*(0, Math.*min*(1,

*f*(1 - **agingInfluence**, (**double**) age / (2 \* **matureAt**))));  
}  
  
*// A function on [0, 1] that intersects (0,h), (0.5, 1), (1,h)***private static double** f(**double** h, **double** x) {  
 **double** a = 4 \* h - 4;  
 **double** b = -a;  
 **double** c = h;  
 **return** a \* x \* x + b \* x + c;  
}

**def** aging: Parametric[Aging] =  
 **for** {  
 matureAge <- *intParam*(**"Maturity Age"**,default = 5,minValue = 0,maxValue = 20)  
 agingInfluence <- *doubleParam*(**"Aging influence"**, 0.3)  
 } **yield new** Aging(matureAge, agingInfluence)

**שאלה 3: מטריקת מרחק והאלגוריתם הMeta Genetic**

מרחק בין גנים ממומש לכל ייצוג בעייה שיש לנו, ע"י הממשק הבא:

**public interface** Metric<A> {  
 **double** distance(A x, A y);  
}

אינווריאנט – מרחק בין גנים יהיה מנורמל לערכים שבין 0 ל-1 – דבר שיקל על האלגוריתמים שמשתמשים במרחק.

המרחק בין גנים הכרחי לשם שיפור הDiversity באופן יותר חכם מאשר רק עם fitness – וזאת בעזרת גיוון אמיתי בין הגנים עצמם.

בכך שאנו נדרוש מרחקים גדולים בין גנים, אך כולם יכולים להיות בעלי fitness מאוד טוב, אנו מאפשרים גידול Local Minima רבים בו זמנית באותו הGeneration, ובכך הם יכולים גם לתרום לפתרון הבעייה בעזרת שילובים של features מכל מינימום לוקאלי.

בנוסף, כך אנו נקבל יותר פתרונות אופטימליים מאשר אחד, מה שהיה קורה אם לא היינו דורשים גיוון בגנים עצמם, וכל הPopulation היה מתכנס לגן מסוג יחיד.

מאוד רואים תופעה זו של שלמות ופתרונות רבים באלגוריתם הMeta Genetic, בחיפוש על פרמטרים הכוללים גם זיהוי והתמודדות עם Local Minimum, במציאת הפרמטר של Gene Similarity Threshold של הבעייה שעליה עושים אופטימיזציה: מהו סף הדימיון שממנו אנו נכנסים למצב התמודדות עם Local Minima.

ברוב הפעמים, האלגוריתם הMeta גנטי ימצא 2 Thresholds אופטימליים לכל בעייה, ועבורם גם התפלגות שונה לשאר הפרמטרים.

בנוסף, שמנו לב שהאלגוריתם הMeta Genetic תמיד התכנס מאוד חזק לפרמטרים הבאים:

Local Optimum: Hyper Mutation: >99% (or a very large value)

Local Optimum: Elitism Rate: <1% (or a very small value)

והדבר מאוד אינטואיטיבי ונכון, מכיוון שאם אנחנו במינימום לוקאלי, אז כדי לצאת ממנו צריך להכניס הרבה אקראיות, ולהקטין את ההיצמדות למינימום הלוקאלי הנוכחי, וזה באמת עובד, כי עבור פרמטים אלה (ושאר הפרמטרים שנבחרו), אנו מקבלים את זמן ההתכנסות האופטימלי והטוב ביותר.

**המטריקות**

על מנת לרשום פונקציות מרחק טובות, או אף 'נכונות' אנו הלכנו לפי הגישה הבאה:

מרחק בין שתי גנים אמור ליצג את אורך ה'מסלול' במרחב בו יש לנו פעולות שינוי מתאימות לבעייה, כך שאנחנו עוברים מגן אחד לשני.

בעיית חיפוש מחרוזת:

מרחק ההאמינג של המחרוזות מנורמל ביחס לאורך המילה.

קוד:

**public static double** distance(**char**[] elem, **char**[] target) {  
 **int** len = Math.*min*(elem.**length**, target.**length**);  
 **int** fitness = 0;  
 **for** (**int** i = 0; i < len; i++) {  
 fitness += *invIndicator*(elem[i] == target[i]);  
 }  
 **return** (**double**) fitness / target.**length**;  
}

בעיית n המלכות:

הסתכלנו על המרחק שבין לוחות כהפרש שבין סכום כל המרחקים שבין מלכות סמוכות על הלוח. מאחר שהייצוג שלנו למלכות על ידי פרמוטציה הינו **אינווריאנטי לשיקוף ואינווריאנטי ל shift אופקי ואנכי**, רצינו מרחק שייצג את הפרמוטציה בהתאם, כך שנתחשב ללוחות שקולים או דומים תחת אינווריאנטים אלה. רצינו ייצוג של ה'דרך' או ה'מסלול' שצריך כדי לעבור מלוח ללוח.

קוד:

**override def** distance(x: QueenPermutation, y: QueenPermutation): Double = {  
 **var** deltas = 0  
 **val** length: Int = x.permutation.length  
 **for** (i <- 0 until length) {  
 **val** delta1 = *abs*(x.permutation(i) - x.permutation((i+1) % length))  
 **val** delta2 = *abs*(y.permutation(i) - y.permutation((i+1) % length))  
 deltas += *abs* (delta1 - delta2)  
 }  
 **val** res = deltas / ((length - 1) \* length)  
 *assert*(res >= 0 && res <= 1)  
 res  
}

בעיית השק:

המרחק בין שני מופעים של בעיית השק יהיה המרחק האוקלידי שבין מערכי כמות ה-items שנלקחו, באופן המנורמל ביחס לשק כולו (הסתכלנו על מערך כמויות כנקודה במרחב הn מימדי). ייצוג כזה של מרחק מבטא כמה יש 'לעבור' מלקיחת מוצרים אחת לאחרת. כמה יש לשנות את מערכי הכמויות שלנו מתשובה אחת לשניה.

קוד:

**override def** distance(x: KnapsackElement, y: KnapsackElement): Double = {  
 **val** instance = x.instance  
 **val** capacity = instance.capacity  
 **def** percentFull(index: Int, amounts: Array[Int]): Double = {  
 **val** maxItems: Double = capacity / instance.items(index).weight  
 amounts(index) / maxItems  
 }  
 **def** normalizedAmounts(amounts: Array[Int]): Array[Double] = {  
 Array.*tabulate*(amounts.length)(i => percentFull(i, amounts))  
 }  
 **val** dist = *arrayDistanceD*(normalizedAmounts(x.amounts), normalizedAmounts(y.amounts))  
 *assert*(dist >= 0 && dist <= 1)  
 dist  
}

בעיית מציאת מינימום של פונקציה:

המרחק האוקלידי שבין הנקודות החשודות כמינימום.

קוד:

**override def** distance(x: FuncSolution, y: FuncSolution): Double = {  
 Distance.*euclidianDistance*(x.xInRange, x.yInRange, y.xInRange, y.yInRange)  
}

**def** euclidianDistance(x1: Double,y1: Double, x2: Double,y2: Double):Double = {  
 *sqrt*(*square*(x1 - x2) + *square*(y1 - y2))  
}

**שאלה 4:**

קריטריון שונות לאבחון מינימום לוקאלי:

כאשר סטיית התקן של ה-fitness של האוכלוסיה קטנה מסף מסויים – אבחנו שאנו נמצאים במינימום לוקאלי.

קוד:

**case class** StdDevLocalOptimaDetector[A](stdDevThreshold : Double) **extends** LocalOptimaSignal[A]{  
 **override def** isInLocalOptima(population: Population[A]): Boolean = {  
 **val** fitnesses = population.*population*.map(\_.*fitness*)  
 **val** (avg, stdDev) = Util.*stdDev*(fitnesses) (*identity*)  
 stdDev < stdDevThreshold  
 }  
}

קריטריון דימיון פרטים לאבחון מינימום לוקאלי:

כאשר סטיית התקן של המרחקים מהפרט הטוב ביותר לשאר קטנה מסף מסויים – אבחנו שאנו נמצאים במינימום לוקאלי.

קוד:

**class** GeneSimilarityLocalOptimaDetector[A](distanceThresh : Double ,metric: Metric[A]) **extends** LocalOptimaSignal[A]{  
 **override def** isInLocalOptima(population: Population[A]): Boolean = {  
 **val** distances = **for** {  
 g1 <- population.*population* } **yield** metric.distance(population.*population*(0).*gene*, g1.*gene*)  
 **val** distanceStdDev = Util.*stdDev*(distances)(*identity*).\_2  
 distanceStdDev < distanceThresh  
 }  
}

**שאלה 5:**

Hyper mutation:

כאשר אובחן שאנו נמצאים במינימום מקומי, נגדיל את הסיכוי למוטציה.

Niching:

כאשר אובחן שאנו נמצאים במינימום מקומי, עבדנו עם פונקציית fitness חדשה שהענישה גנים הדומים לגנים אחרים.

קוד:

**public double** mapFitness(Population<A> population, Gene<A> gene) {  
 **double** sumOfSharingFunc = 0;  
 **for** (**int** i = 0; i < population.**population**.**length**; i++){  
 **double** distance = **metric**.distance(

gene.**gene**,

population.**population**[i].**gene**);  
 **if** (distance < **sigmaShare**)  
 sumOfSharingFunc += 1 - Math.*pow*(distance / **sigmaShare**, **alpha**);  
 }  
 **return** 1 - ((1 - gene.**fitness**) / sumOfSharingFunc);  
}

Random immigrants:

כאשר אובחן שאנו נמצאים במינימום מקומי, נוסיף לאוכלוסיה שלנו מספר גנים רנדומיים חדשים וטריים שירעננו את מאגר הגנים שלנו.

קוד:

מיד אחרי האליטיזם יבואו המהגרים הרנדומיים (במידת הצורך)

**int** popSize = population.**population**.**length**;  
**int** elites = (**int**) (popSize \* **elitismRate**);  
*elitism*(population, buffer, elites);  
  
**int** immigrants = (**int**) **randomImmigrantsPercent** \* popSize;  
**for** (**int** i = elites; i < (elites + immigrants) && i < popSize; i++) {  
 buffer.**population**[i].**gene** = alg.randomElement(rand);  
}

**שאלה 6:**

**שאלה 7:**

חזית פרטו אופטימל שמצאנו עבור הפונקציות המופיעות להלן הינה

קוד גנטי עבור מציאת הפרטו פרונט:

**חלק ב – אפקט בולדווין**

\*\*\*\*\*\*\*\*\*שאילן יחפור על זה כאן... ליובל אין כוח\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*

לפי הניסויים שעשינו אפקט בולדווין כן קורה בלה בלה בלה

קוד:

**override def** fitness(gene: Array[BaldwinBit]): Double = {  
 **val** remainingItarations = BaldwinBitString.*localSearchToBest*(gene, maxIterations, isFinal, rand)  
 **val** rawFitness = 1 + 19.0 \* remainingItarations.toDouble / maxIterations  
 **val** normalized = (20 - rawFitness) / 20  
 *assert*(normalized >= 0 && normalized <= 1)  
 normalized  
}

גרפי הרצה:

לפי מספר הרצות שעשינו, ראינו שהחל משלב מסויים האחוז של ממוצע הביטים הנכונים בכל דור גדל בצורה ניכרת. ההסבר לכך: אפקט בולדווין.

אחוז ממוצע של ביטים נכונים בכל דור:

אחוז ממוצע של ביטים לא נכונים בכל דור:

בהתחלה מתחילים עם 25%, שמתחיל מעט יציב, ואז הולך ויורד עד לכמעט התאפסות. פה ושם יש גנים עם מעט לא נכונים עקב מוטציות וזיווגים

אחוז ממוצע של למידה של ביטים:

כפי שניתן לראות, יותר ביטים נלמדים מאשר לא נלמדים. סך הכל, כמות סימני השאלה הולכת ופוחתת ההסבר לכך: אפקט בולדווין.