# שיערוך תמונה בעזרת שיטות היוריסטיות

## דו"ח סיכום במסגרת הכשרת צוות אלגוריתמיקה מבצעית

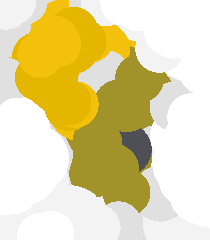
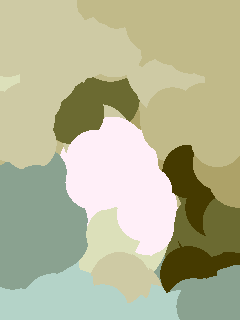
### תיאור הבעיה

נתונה תמונה (חתול חמוד, הומר סימפסון חמוד וכדו').



חתול והומר חמודים

בעזרת ציור מאה עיגולים על קנבס ריק יש לקרב את התמונה "עד כמה שאפשר", בעזרת שתי השיטות (או וריאציות שלהן): Hill Climbing, Simulated Annealing.



שיערוך בעזרת מאה עיגולים

#### מבט (יותר) פורמלי

* בעיית החיפוש מתחילה ממצב התחלתי של קנבס ריק.
* כל ציור על הקנבס מהווה מצב.
* המצב האופטימלי הוא המצב של התמונה המקורית.
* ערכו של מצב נמדד על ידי המרחק בין התמונה של אותו המצב לבין התמונה המקורית (מרחק יותר קטן מניב ערך יותר גדול), כאשר ניתן להגדיר מרחק בכמה צורות.
* ניתן להגדיר מעברים חוקיים בין מצבים באופנים שונים. אנחנו נגדיר שניתן לעבור מתמונה מסויימת רק לכל התמונות שזהות לה ונוסף להן רק עיגול אחד.
* המטרה היא להגיע למצב המייצג תמונה שמרחקה מהתמונה המקורית קטן ביותר (מטרות משנה – לבחון גם תמונות שמרחקן מהתמונה המקורית גדול יותר, אך הן "לוכדות" תכונות מעניינות של התמונה המקורית)

### תיאור השיטות

#### טיפוס הרים

בחיפוש מסוג "טיפוס הרים", ההתקדמות במרחב המצבים שבעולם הבעיה נעשה תמיד אל הצעד הבא שבו הרווח (המשוייך לכל צעד) ממוקסם. בהתאם, ההתקדמות נעצרת כאשר מגיעים אל "גבעה", כלומר מצב שערכו גבוה מערך כל המצבים שסביבו.

הגישה של שיטת החיפוש הזו חמדנית במובן זה שבכל שלב נעשית הבחירה המניבה את הרווח המיידי הגדול ביותר, וכפי שנוטה לקרות בשימוש בגישות חמדניות, אלגוריתם החיפוש עלול להיתקע בנקל במינימום מקומי.

ישנן גרסאות נוספות לאלגוריתם שמנסות להתגבר על הבעיה הנ"ל, ביניהן:

1. טיפוס הרים סטוכסטי – בחירת הצעד הבא בצורה אקראית, כאשר ההסתברות לבחירת צעד נקבעת לפי רמת העלייה בצעד הזה (הפרש בין ערך הצעד הבא לערך הצעד הנוכחי).
2. טיפוס הרים - בחירה ראשונה – הגרלת הצעד הבא עד שמתקבל צעד בעל ערך גבוה משל הצעד הנוכחי.
3. טיפוס הרים עם התחלה רנדומית – הרצה שוב ושוב של טיפוס הרים מנקודות התחלה שונות, ובחירת הפתרון הטוב מביניהם.

#### חישול מדומה

כאמור, טיפוס הרים לעולם לא בוחר בצעד שבמורד הגבעה, אלא לכל היותר בוחר צעד לא אופטימלי במעלה הגבעה. בבירור, גישה כזו עשויה להיתקע במקסימום מקומי (בגבעה עליה החיפוש התחיל) כך שהפתרון אינו "שלם" (אינו בהכרח מוצא את הפתרון – האופטימום של בעיית האופטימיזציה).

בניגוד לכך, בחירת צעד אקראי לחלוטין, למשל בהתפלגות אחידה, היא שיטה "שלמה" שמובילה בסופו של דבר לפתרון האופטימלי, אלא שהיא עושה זאת באופן לא יעיל בעליל.

גישת חישול מדומה, במובן מסויים, לוקחת את הטוב שבשתי הגישות. בשיטה זו בוחרים בכל שלב צעד רנדומי, ומקבלים אותו אם ערכו גדול מהערך הנוכחי. אחרת, כלומר, אם הצעד המוגרל מפחית מהערך הנוכחי, הוא מתקבל בהסתברות קטנה מ1. ההסתברות קטנה באופן מעריכי ככל שהצעד יותר "גרוע", כלומר, ככל שהערך בבחירה בו קטן, ובנוסף ההסתברות קטנה ככל שמספר האיטרציות הולך ועולה.

האינטואיציה בגישת החישול המדומה היא שעפ"ר התנועה תהיה בטיפוס למעלה, אלא שישנה גם אופציה לבחור בירידה מהגבעה כדי להתחמק ממינימום מקומי. באופן הגיוני, ככל שהצעד יותר גרוע ההסתברות לבחירתו יותר קטנה, ובנוסף, ככל שקרובים לתחילת החיפוש ניתנת יותר אפשרות וחופש לבחור באופציות הפחות טובות.

#### סיכום

התמצית של הגישות, כפי שאנו רואים אותה, ובפרט בהקשר של שיערוך תמונות, היא כדלהלן.

* שיטת טיפוס הרים בוחרת תמיד בצעדים שרק משפרים את ערך המצב.
* שיטת חישול מדומה מאפשרת בהסתברות מסויימת צעדים שמורידים הערך כדי לאפשר התנסות בכיוונים שונים שייתכן שיאפשרו הישגים שלא היו מתקבלים במינימום המקומי אליו שיטת טיפוס הרים הייתה מובילה.

### שיטות פעולה

בחלק זה נציג את נקודות המפתח בהגדרת ופתירת הבעיה, ואת הגישות השונות שניסינו.

#### בחירת צבעים

ראשית יש להבין מהו מרחב הצבעים שבו נשתמש לציור העיגולים.

* באופן נאיבי אפשר להתבונן בכל המרחב של RGB, הכולל אפשרויות, אבל די ברור שמדובר באפשרות לא מוצלחת משתי סיבות. ראשית, מרחב המצבים נעשה כל כך גדול עד כדי שלא ניתן לבצע יותר ממספר מועט של איטרציות באלגוריתם החיפוש בזמן סביר. יתר על כן, בכל תמונה סבירה ישנו מספר מועט של צבעים שמכסים את כולה, או כמעט את כולה, וכל נסיון לחיפוש בצבעים אחרים הוא נסיון מיותר.
* אפשר להתבונן בתת קבוצה של צבעים מגודל מצומצם, אך ברור שככל שנקטין את הקבוצה, אם לא נתאים את הצבעים במיוחד לתמונה המסויימת, נחווה ירידה משמעותית בטיב התמונה שניצור.
* אפשרות אחרת היא לבצע שלב מקדים של דגימת צבעים מהתמונה לשם קביעת מרחב צבעים מצומצם ויעיל שבו האלגוריתם יבצע את החיפוש. הדגימה יכולה להתבצע באופנים מגוונים. נציע אפשרות אחת, אך לא קשה לחשוב על וריאנטים נוספים.

קובעים היפר פרמטר אחד המייצג את מספר הצבעים הרצוי במרחב הצבעים, וקובעים היפר פרמטר נוסף המייצג את המרחק המינימלי הרצוי בין הצבעים (כדי שלא ידגמו שני צבעים דומים, כמו שני גוונים כמעט זהים של לבן למשל).

כעת מבצעים דגימה יוניפורמית של פיקסל מהתמונה ומוסיפים את הצבע באותו פיקסל לרשימה המייצגת את מרחב הצבעים, כל עוד אין ברשימה את מספר הצבעים הרצוי וכל עוד הצבע החדש שונה מספיק משאר הצבעים הנמצאים ברשימה.

נשים לב שהדגימה היוניפורמית מעל הפיקסלים מהווה למעשה דגימה לא אחידה של הצבעים, שכן צבע שמופיע יותר בתמונה בעל הסתברות יותר גבוהה להידגם.

* אפשרות נוספת היא שלאחר שנקבע מרכז של עיגול אותו נרצה לצייר, נקבע את צבע העיגול להיות כמו צבע המרכז. נשים לב ששיטה זו נותנת יותר הזדמנות לצבעים פחות שכיחים להופיע. לדוגמא, הצבע האדום בפה של הומר סימפסון, צבע פחות שכיח בתמונה, הופיע יותר בשימוש בשיטה זו.

יתרון משמעותי של האפשרות הזו לעומת הקודמת היא זמן ריצה מהיר פי כמה (שכן צבע העיגול נקבל מיידית ולא צריך לעבור על כל הצבעים במרחב הצבעים). מאפיין נוסף של השיטה הזו לעומת הקודמת, כאמור, הוא מתן הזדמנות לצבעים פחות שכיחים להופיע (גם אם הצבעים הפחות שכיחים יורידו את ערך התמונה).

* עוד אפשרות היא לקבוע את צבע העיגול לפי צבע רוב הפיקסלים של העיגול הנקבע על פי המרכז והרדיוס. לא ראינו יתרון בשיטה זו לעומת שתי השיטות הקודמות, לא מבחינת התוצר ולא מבחינת זמן הריצה.

#### קביעת ציון למצב

שאלה נוספת שצריך להתייחס אליה היא כיצד למדוד את המרחק בין התמונה הנוכחית לתמונה האופטימלית.

* מדד הגיוני ואינטואיטיבי שמקובל להשתמש בו הוא הmean square error שמודד את ריבועי ההפרשים בין הפיקסלים של התמונות. נשים לב שההעלאה בריבוע מאפשרת להתייחס להפרשים חיוביים ושליליים אותו הדבר.
* ניסינו מדד נוסף בין קידודים של תמונות ע"י מתודה של ספרייה שמקודדת תמונות, אבל התוצאות שהתקבלו היו מאוד לא טובות. לא נכנסנו לפרטים של הספרייה, אך ככיוון עתידי הדבר עשוי להיות נכון.

#### מיקום המעגלים

* מיקום המעגלים בכל איטרציה נקבע בצורה רנדומית בבחירה יוניפורמית.
* אפשרויות נוספות שניתן היה לעשות הן להקצות סך מסויים של המעגלים לאיזורים שונים בתמונה. למשל להקצות מספר מועט של מעגלים לסביבה שקרובה יותר למסגרת, ולהקצות יותר מעגלים (עם גדלים מתאימים) לאזור המרכזי, בהנחה סבירה שהוא יותר מרובה בפרטים (בתמונות עליהן עבדנו בודאי, אבל כנראה הדבר נכון גם להתפלגות התמונות בעולם).
* אופציה נוספת היא להגריל בכל איטרציה הרבה נקודות ולקחת את הטובה מביניהן (או את אחת האפשרויות הטובות, אפילו אם היא לא האופטימלית). החסרון באפשרות הזו הוא הכבדה בזמן הריצה. ניתן באופן דומה לקחת את האפשרות הטובה מבין כמה אפשרויות במקומות אחרים בpipeline של האלגוריתם.

#### רדיוס המעגלים

* רדיוס המעגלים מתחיל מערך חצי מרוחב התמונה ודועך עם מספר האיטרציות. מימשנו את הדעיכה על ידי חלוקה בלוגריתם של מספר האיטרציות (ועוד 2, כדי להימנע מחלוקה ב0).
* ניתן לחשוב על אופנים נוספים להגדיר את הדעיכה. אפשר לחלק במספר האיטרציות ולא בלוגריתם, אלא שאז הדעיכה מהירה במיוחד ותחת הגבלה של מאה עיגולים לא מתקבלת תמונה טובה.
* הרדיוס גם יכול להיות מוגרל, למשל בהגרלה יוניפורמית בתחום שבין לוגריתם מספר האיטרציות לחצי רוחב התמונה. התמונות שהתקבלו היו פחות טובות. לא קשה לחשוב על וריאנטים נוספים דומים.

#### מימוש טיפוס ההרים

פרטי המימוש האלגוריתמיים של שיטת טיפוס הרים הם כדלהלן.

בכל איטרציה יצרנו קבוצות של מצבים שכנים למצב הנוכחי, כאשר מספר הקבוצות הוא היפר פרמטר.

אופן יצירת קבוצת שכנים משותף למימוש של שיטת חישול מדומה. נציג אותו בהמשך.

מתוך כל קבוצות השכנים שנוצרו, נבחר השכן בעל הערך הגדול ביותר. אם ערכו גדול מערך המצב הנוכחי האלגוריתם מתקדם אליו. אחרת, האלגוריתם חוזר על האיטרציה.

האלגוריתם מורץ מספר פעמים (גם כן היפר פרמטר), ומתוך כל התמונות שנוצרו נלקחת התמונה הקרובה ביותר לתמונה האמיתית.

* באופן כללי ניתן לסכם ולומר שהשכבה של טיפוס ההרים באלגוריתם שלנו חמדנית לחלוטין. מה שמאזן את החמדנות הזו הוא המימוש הפנימי של יצירת קבוצת השכנים שמכיל אקראיות רבה, וכן לקיחת התוצר הטוב ביותר מתוך כמה הרצות של האלגוריתם.

#### מימוש חישול מדומה

בכל איטרציה של שיטת חישול מדומה אנחנו פועלים באופן אחד מתוך שניים: exploration or exploitation, על ידי בחירה הסתברותית הנשלטת על ידי היפר פרמטר וטמפרטורה.

הטמפרטורה דועכת עם מספר האיטרציות (על ידי חלוקה בלוגריתם מספר האיטרציות), כדי לאפשר יותר "מחקר" בשלבים הראשונים ויותר "פיתוח" בשלבים האחרונים של החיפוש.

פעולה באופן של explotation זהה לאופן הפעולה של טיפוס הרים.

לעומת זאת, בפעולה באופן של explotation אנחנו דוגמים שכן מבין קבוצת השכנים המתקבלת, כאשר ההתפלגות נוצרת על ידי הציונים של השכנים ופונקציית softmax. כך אנחנו מאפשרים למהלכים פחות טובים להתקבל, אך בהסתברות הולכת ופוחתת לפי ערכם השלילי.

* באופן כללי ניתן לסכם ולומר שהמימוש מכיל בתוכו את העקרונות של שיטת חישול מדומה המקורית. ניתנת הזדמנות גם לאפשרויות פחות טובות וירידה מההר, אך ההסתברות שהן תיקרנה הולכת ופוחתת לפי טיב האפשרויות ולפי התקדמות האלגוריתם.

#### מימוש יצירת קבוצת שכנים

יצירת קבוצת השכנים משותפת לשתי השיטות, טיפוס הרים וחישול מדומה.

יצירת קבוצת השכנים נעשית בהתאם לדרכי המימוש שתוארו לעיל.

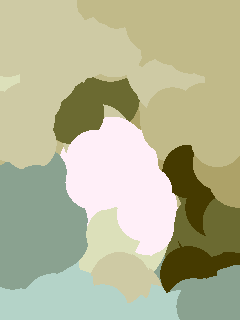
בוחרים רדיוס שמשותף לכל השכנים שבקבוצה (כפונקציה דטרמניסטית של מספר האיטרציה או בהגרלה מתוך תחום מסויים) ובוחרים מרכז אקראי.

כעת מוסיפים את כל התמונות עם העיגולים מכל צבע מתוך מרחב הצבעים בעלי המרכז והרדיוס האלו, וזו קבוצת השכנים.

ניתן לחשוב על וריאציות שונות, כגון הוספת שכנים שונים עם רדיוסים שונים או מרכזים שונים. הדבר מוסיף לזמן הריצה, ומבחינה תיאורטית שקול לדברים האקראיים שהכנסנו בנקודות אחרות באלגוריתם.

### תוצרים

בחלק זה נציג תמונות משוערכות שונות שקיבלנו, נציג את התנאים בהם הן התקבלו וננסה להסביר את המאפיינים שלהן בהתאם לשיטות הפעולה המוצגות לעיל.

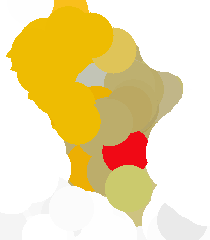




תמונות של החתול החמוד שהתקבלו בשיטת טיפוס הרים על ידי רדיוס קבוע (מימין) ורדיוס מוגרל (משמאל)

ישנו הבדל בין הצבעים בתמונות מכיוון שמרחב הצבעים בכל תמונה שונה במקצת.

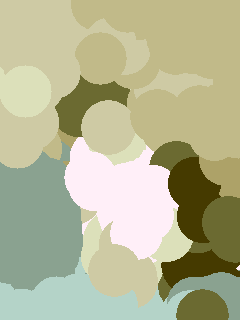
בתמונה השמאלית גדלי הרדיוסים הוגרלו, ובמקרה זה התקבלה צורה פחות מסודרת מאשר הצורה שבצד ימין שנוצרה מרדיוסים שהלכו וקטנו בצורה דטרמניסטית.

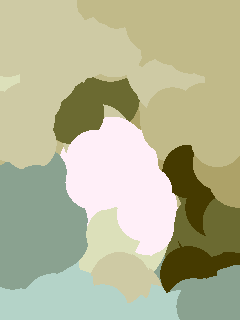




תמונות של הומר החמוד שהתקבלו בשיטת טיפוס הרים על ידי שימוש במרחב צבעים (מימין) ושימוש בצבע המרכז (משמאל)

נדמה כי השימוש במרחב צבעים עשוי להניב תוצאה שמביאה למרחק קטן יותר בין התמונה המתקבלת לבין התמונה האמיתית, שכן מרחב הצבעים תופס את הצבעים השולטים בתמונה ומתעלם מצבעים שלא נוטים להופיע הרבה, אך בכל זאת ניתן לראות בתמונה השמאלית שהצבע הזניח – אדום הצליח להופיע בתור הפה של הומר. אולי מבחינת ציון העיגול האדום מפחית את הערך, אך כתוצר למשתמש הדבר גורם לפליטת דופמינים.

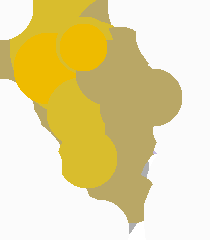


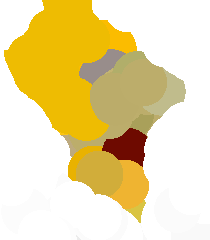


חתול חמוד שנוצר על ידי טיפוס הרים (ימין) לעומת חתול חמוד שנוצר על ידי חישול מדומה (משמאל)

ניתן לראות שהחתול שנוצר על ידי בחירות חמדניות נראה יותר מדוייק גם מבחינת קווי מתאר וגם מבחינת צבעים מאשר החתול שנוצר תוך כדי מתן הזדמנויות לבחירות פחות טובות.

אכן, נדמה כי במשימת האופטימיזיציה הספציפית הזו, מכיוון שעל ידי הוספת עיגול אפשר לתקן את המצבים הקודמים ואין באמת אפשרות להיתקע במינימום מקומי כמו במשימת טיפוס הרים אמיתית, כנראה האפשרות של טיפוס הרים (על אחת הוריאנטים שלה) תמיד תהיה יותר טובה.

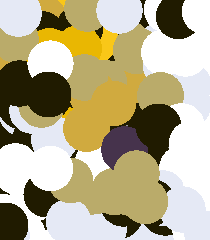


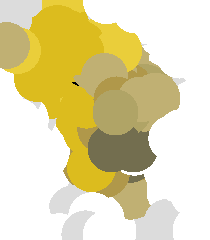


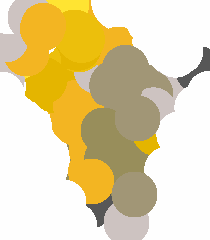
הומר חמוד שנוצר על ידי ציון MSE (מימין) לעומת הומר חמוד שנוצר על ידי השוואת קידודים (משמאל)

ניתן לראות כי השימוש בריבועי ההפרשים מכמת בצורה יותר טובה את רמת הדמיון בין התמונה לתמונה האופטימלית מאשר הפרש הקידודים, הגם שהפרש הקידודים גם כן משרה איזשהו סדר הגיוני בין התמונות.



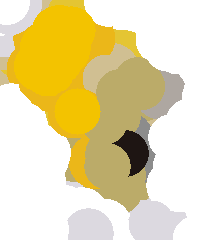


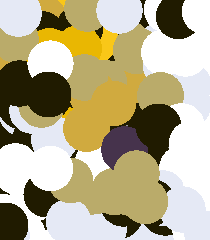




הומר החמוד שנוצר על ידי חישול מדומה בלי טמפרטורה עם הסתברות לexploitation של 0.25,0.5,0.75 ו0.9

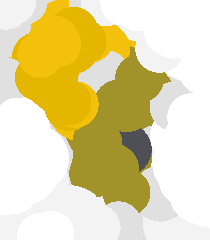
בלי טמפרטורה ניתן לראות הבדלים מובהקים בין ההסתברויות השונות הניתנות ליחס בין exploration and exploitation. ההסתברויות שניתנו בתמונות, לפי סדר מימין לשמאל ומלמעלה למטה, הן: 0.25, 0.5, 0.75 ו0.9. רואים שיפור משמעותי ככל שמנסים יותר "לפתח" ופחות "לחקור", כאמור, ללא האלמנט של הטמפרטורה שמאפשר יותר לחקור באיטרציות הראשונות ופחות באיטרציות האחרונות.





הומר החמוד עם הסתברות exploitation 0.25 בלי טמפרטורה (מימין) ועם טמפרטורה (משמאל)

כאשר מכניסים את אלמנט הטמפרטורה, רואים בבירור שאי הסדר שנגרם בגלל עודף המחקר הופחת.



הומר החמוד עם הסתברות exploitation 0.5 בלי טמפרטורה (מימין) ועם טמפרטורה (משמאל)

ההבדלים ברורים גם כאשר הסתברות המחקר יותר נמוכה.