UTF-8:

קידוד שנוצר על יד רוב פייק וקן תומפסון וניתן לקודד בו כל תו על ידי שימוש באחד עד 4 בתים

קידוד זה מוסיף חסכון בזכרון עמידות בפני איבוד או השחתת בתים ותיאמות לאחור לאסקי

ו מורה לנו איך לקחת את המספר הייצוגי ולשמור אותו במחשב לדוגמא אם ברשותנו טקסט

שאנחנו מעוניינים להכניס אותו למחשב מחליפים כל תו למספרו הייצוגי ונקבל רצף של מספרים .

ונשמור אותם במחשב .הקידוד הוא תלוי תו ולכל תו יש מספר שמייצג אותו .וכל מספר מייצג תו אחד בלבד .נהוג לרשום את המספרים שמייצגים את התווים דרך מספריהם הייצוגיים ולא דרך השמות .

הקידוד נעשה באמצעות מספר משתנה של בתים ורצפי הסיביות .על פי חוקי הקידוד נדרש שיתקיים :

הסיבית המשמעותי ביותר בבית המייצג תו אחד היא 0

הסיבית המשמעותי ביותר בבית שהוא חלק ממספר בתים שמייצגים תו אחד היא 1

כל בית שפותח רצף של בתים שמגדירים תו בודד מתחיל ב 1 ואז 0 לדוגמא אם תו מיוצג על ידי שני בתים הוא יתחיל ב110 ואם על ידי 3 אז 1110

כל בית שהוא בית נוסף ברצף בתים שמייצגים תו בודד מתחיל ב10

בזכות חוקים אלו לא נוכל להתבלבל בין בית שמייצג תו האסקי לבית שהוא חלק מכמה בתים.

יתרונות וחסרונות של הקידוד:

מכיל בתוכו את טבלת אסקי .כל טקסט שקודד באסקי תקני גם לפי הקידוד שלנו ולכן אין צורך בשום פעולת המרה לטקסטים קיימים שכבר קודדו באסקי

ניתן להפעיל כל אלגוריתם חיפוש שמבוסס על חיפוש לפי בתים על טקסטים שקודדו .

ניתן לזהות בצורה יחסית פשוטה טקסטים שקודדו לפי הקידוד הנ"ל

חסרונות :

מפענח שנכתב בצורה לא טובה ולא תואם גרסאות עדכניות של סטנדרטים יכול לקבל בטעות גרסאות שונות של ייצוג תווים ולפענח אותם לתווי יוניקוד שהם כביכול מייצגים כך מידע יכול לדלוף מעבר להליכי בדיקה ואימות שונים .