

Repository of logically consistent real-world Boolean network models

רקע

במאמר זה נתאר מאגר חופשי של יותר מ- 210 רשתות בוליאניות שזמינות במגוון רחב של פורמטים. במאגר יש סט תהליכים שמוודאים שאכן המודלים נכונים לוגית. סט התהליכים הנ"ל עוזר לתקן יותר מ- 400 בעיות פוטנציאליות במגוון רחב של מודלים נפוצים.

המאמר נוצר בעקבות התאוצה המצטברת בעולם הניתוח חישובי על בסיס מאגרי מידע, מה שמאפשר עידן חדש של אטומיזציה במערכות ביולוגיות, וכן לפקח בצורה איכותית ולהשוות בצורה טובה בין מודלים.

רקע תיאורטי

רשתות בוליאניות הן אלמנט בסיסי במידול מערכות ביולוגיות, כאשר בדרך כלל הן ממודלות ומנותחות באופן ידני. עם המעבר לניתוח באופן אוטומטי צצו כמה מגבלות כמו חוסר עקביות במידול המודלים והגבלה רק לידע שנצפה במאגר המידע, ולא בהכרח מכסה את כל המקרים הקיימים. מכאן מגיע הצורך במאגר BBM שמכסה כמות מכובדת של מודלים ממאגרים שונים ובפורמטים שונים, תוך שמירה על נכונות לוגית של המודלים.

אלמנטים של מודל רשת ביולוגית

רשתות בוליאניות הן אוסף של פונקציות עדכון בוליאניות מעל משתנים, כאשר כל פונקציה תלויה באיזושהי תת קבוצה של משתנים. הרשתות מיוצגות ע"י גרף מכוון עם תוויות, שמייצגות תלויות בין משתנים. הדרישות לחיוניות ומונוטוניות, דואגות לכך שהפונקציות יעמדו בהנחות של הביולוגיה ובכך מייצגת הרשת הביולוגית.

הוספה של מודל חדש

התהליך של הוספה של מודל חדש למאגר, נעשה ברובו באופן ידני שמוודא שמדובר במודל שממדל מערכת ביולוגית אמיתית. זה יכול לבוא ממאגר אחר מוכר, פרסומים של מודלים וכו'. המאגר מכיל גם רשתות לא בוליאניות, אך כדי למקסם יעילות והתאמה עם הכלים הקיימים, כל מודל כזה עובר לתצורה בוליאנית.

במאגר לכל מודל יש מזהה ייחודי ושם קריא, ביחד עם קבצי קלט תקינים באחד מכמה פורמטים מקובלים. אם המודל לא זמין באחד מהפורמטים האלה, המודל משונה באופן ידני כדי להתאים תוך שמירה על ההצגה המקורית שלו.

הפיפ ליין מתרגם את הקלט לפורמטים, ומבצע נורמליזציה כדי לוודא התאמה עם הכלים. זה כולל לתת שמות ייחודיים למשתנים ולנרמל את משתני הקלט, שזה בעצם משתנים בלי רגולציות. הפיפ ליין מייצא את הקלטים לאחר הנרמול ומאפשר למשתמש לבחור את הייצוג שישמש אותו בצורה הטובה ביותר.

הפיפ ליין גם מוודא את הגרף אל מול פונקציות העדכון בעזרת BDD. אם נמצאו אי עקביות אז משנים באופן ידני את הגרף ומתעדים את השינוי. שגיאות נפוצות הן בעיקר כשחסר משתני קלט, אי התאמה בתוויות וכו'. בנוסף הפיפ ליין מסיר רכיבים שלא בשימוש ושומר על גרף קשיר חלש.

המאגר מנוהל דרך גיט, כל מודל זמין בשלושת הפורמטים: bnet, aeon, sbml יחד עם קבצים שמתארים את המקור של המודל, שינויים ומילות מפתח. בסביבה ניתן בצורה נוחה לייצא מהודרות מתואמות אישית של המאגר, לבחור פורמט, ייצוג קלט וסינון על פי מאפיינים מבניים.