**מבנה המחשב – פרוייקט ISA, דוקומנטציה**

איתי מרקמן – 206900193

עידו גפן

בן דב בלוך

**רקע קצר**

בפרוייקט אנו נדרשים לחקור את עבודתו של מעבר בשם SIMP הדומה למאבד MIPS ולראות איך הוא מריץ תוכנות הכתובות בשפת האסמבלי הייחודית שלו.

בשפה זו ישנם 19 פקודות אסמבלי המוצגות להלן, כאשר בקובץ הזיכרון(memin.txt, memout.txt) כל פקודה רשומה כשורה הקסאדצימלית ברוחב 8 תווים. או 32 ביטים. הפקודה בנויה ככה מימין לשמאל:

תווים 1:3 – שדה ה-Immediate שיוכנס לתוך אוגר ה-Immediate בביצוע הפעולה

תו 4 – אוגר rt

תו 5 – אוגר rs

תו 6 – אוגר rt

תווים 7-8 –opcode, שם הפקודה

להלן טבלה עם כל הפקודות שתוכנות האסמבלר והסימולטור תומכות בהן:

|  |  |  |
| --- | --- | --- |
| מספר opcode בהקסאדצימלי | שם בשפת אסמבלי | משמעות |
| 00 | Add |  |
| 01 | Sub |  |
| 02 | And |  |
| 03 | or |  |
| 04 | Sll |  |
| 05 | Sra | הזזה אריתמטית עם הארכת סימן |
| 06 | Srl | *הזזה לוגית* |
| 07 | beq |  |
| 08 | bne |  |
| 09 | blt |  |
| 0A | bgt |  |
| 0B | Ble |  |
| 0C | bge |  |
| 0D | jal |  |
| 0E | lw |  |
| 0F | sw |  |
| 11 | reti |  |
| 12 | in |  |
| 13 | out |  |
| 14 | halt | Halt execution – exit simulator |

בכל פקודות הbranch. ההתניה נעשית רק על 11 הביטים הימניים ביותר של האוגרים. ו-PC פירושו מספר השורה הנוכחית וmem זה כתובת של שורה בזיכרון.

הוראות בקבצי אסמבלי מתחילות ב# וישנה אפשרות להכניס מספר של תא בזיכרון כlabel. כלומר לתת לו כינוי ואז לעשות נקודותיים. כך אפשר לקפוץ ממקום למקום בקוד יותר בקלות

פקודת אסמבלי אחרונה היא הפקודה ".word" פקודה זו מאחסנת מילה של 32 ביטים(8 תווים הקסאדצימלים) ישירות בזיכרון. היא בנויה כך:

.word address data

כמו כן למעבד 16 אוגרים עם תפקידים שונים

|  |  |  |
| --- | --- | --- |
| מספר(הקסאדצימלי) | שם | תפקיד |
| 0 |  | תמיד מכיל אפס |
| 1 |  | הפניה לשדה ה-immediate של הפקודה |
| 2 |  | ערך מוחזר מפונקציה |
| 3 |  | ארגומנט של פונקציה |
| 4 |  | ארגומנט של פונקציה |
| 5 |  | משתנה זמני |
| 6 |  | משתנה זמני |
| 7 |  | משתנה זמני |
| 8 |  | משתנה זמני |
| 9 |  | משתנה זמני נשמר |
| A |  | משתנה זמני נשמר |
| B |  | משתנה זמני נשמר |
| C |  | Global pointer(static data) |
| D |  | מצביע למחסנית |
| E |  | Frame pointer |
| F |  | כתובת חזרה לאחר שימוש בפונקציה |

אנו מניחים שתדר המעבד הוא 256 הרץ ושכל פקודה לוקחת מחזור שעון אחד

**האסמבלר –Assembler.c**

תפקיד תוכנת האסמבלר היא לקחת קובץ asm הכולל בתוכו תוכנית אסמבלי בשפה שתוארה בפרק הקודם. ולהמירו לקובץ txt המכיל שפת מכונה הקסאדצימלית המתארת את תמונת הזיכרון לפי ריצת תכנה. האסמבלר בנוי כך שיוכל לתמוך בתוכנות גדולות ככל האפשר(אינסופיות בתאוריה) אך מניח כי לייבל מוכב מ50 תווים. אורך שורת קוד הוא עד 500 תווים, מספר immediate הוא עד 50 תווים, ושם אוגר/אופקוד עד 6 תווים וכתוב בשפה c מתחילתו ועד סופו. ישנו משתנה גלובלי MAX\_LINE המוגדר לעד 500 תווים

**הערה לבודק**– האסמבלר מסתמך על הספריות stdio.h,stdlib.h,string.h ולכן יש לוודא שהן מותקנות לפי הרצה. השימוש בספריות אלו חוקי לפי הוראות הפרוייקט

מבנה עזר –label:

על מנת לתמוך בתוכנות בגודל אינסופי, האסמבלר מאחסן מידע על הקוד שהוא קורא וממיר ברשימות מקושרות הבנויות ממבנים בשפת סי. בכל מבנה מידע על האיבר הנוכחי ומצביע לבא ברשימה. מתחת לכל הגדרה של מבנה בקוד ישנן פונקציות עזר הקשורות אליו. מבנה קוד זה הוא הדבר הקרוב ביותר ל"קוד מונחה עצמים" שניתן לבנות בשפת סי

נסתכל על המבנה label:

// the label linked list will help us store info about the labels. it will be built at the first iteration and used and the second. then destroyed

typedefstructLabel

{

// name will store the name of the label

char name[50];

// location will store the line number and will be the immediate value in related jump and branch commands

int location;

// a pointer to the next label

structLabel\* next;

} label;

הפונקציה המגדירה את המבנה, שכולל מחרוזת nameעד 50 תווים עם שם הלייבל, מספר locationהמתאר את מיקומו בזיכרון(השורה שיקפצו אליה) ומצביע nextללייבל הבא ברשימה

// this function creates a label from the given name and location

label\* create\_label(charname[50], intlocation)

{

// allocate memory for the label and create a pointer to it

label\* new\_label = (label\*)malloc(sizeof(label));

// if allocation successful. insert data to label

if (new\_label != NULL) {

// use strcpy to insert a string

strcpy(new\_label->name,name);

// the other insertions are easy

new\_label->location = location;

new\_label->next = NULL;

}

return new\_label;

}

ה"constructor" של מבנה זה. לוקח מחרוזת ומבנה, מקצה זיכרון, ואם היא מצליחה להקצות זכרון, שמה את מה שהיא קבלה בתאים הרלוונטים ומחזירה מצביע לlabel החדש

// adds a label to the front of an existing label list with given name and location.

// will be used in the first iteration

label\* add(label\* head, charname[50], intlocation) {

// build the label and check for success. otherwise return a null

label\* new\_label = create\_label(name, location);\_ if (new\_label == NULL)

returnNULL;

// the next pointer will point to the original head

new\_label->next = head;

return new\_label;

}

פונקציה זו מקבלת את שם הלייבל ומיקומו ומוסיפה לייבל חדש בתחילת הרשימה.

// this will scan the label list "head" and look for "name". it will return it's location.

// this will be used in the second iteration

int find(label\* head, charname[50]) {

// current - the current label's name

char current[50];

// start with the head

strcpy(current, head->name);

// strcmp returns 0 if names are equal

while (strcmp(current, name) != 0) {

// go to next label

head = head->next;

// in case not found - return -1

if (head == NULL) {

return -1;

}

// update name to current name

strcpy(current, head->name);

}

// return the current's location

returnhead->location;

}

פונקציה זו נקראת בין האיטרציה השנייה של האסמבלר לכתיבת תמונת הזיכרון. היא מקבלת שם של לייבל ומצביע לרשימה. ועוברת על הרשימה איבר איבר(המצביע current מצביע לאיבר הנוכחי) עד שהיא מוצאת את הלייבל שהאסמבלר רוצה לחפש ומחזירה את מיקומו בזיכרון. אם הלייבל לא נמצא מוחזר -1

// destroys the list and frees the memory

void destroy(label\* head)

{

// temp - a pointer to a label we are going to destroy after updating head

label\* temp;

// all the way to the end

while (head != NULL)

{

// temp gets the current node while head advances

temp = head;

head = head->next;

// we destroy temp and free the memory

free(temp);

}

}

פונקציה הנקראת בסוף הקוד שהורסת את הלייבל ומשחררת את הזיכרון.המשתנה temp עוזר בשחרור כך שתמיד יהיה לנו מצביע לראש הרשימה הנוכחי

מבנה עזר - שורת זיכרון MemoryLine:

מבנה זה מתאר שורה בזיכרון. ובו מחרוזות opcode, rd, rs, rt,imm המתארים כל שדה של הפקודה באסמבלי, מספר pos המתאר את מספר השורה בזיכרון אליה צריך לפנות ומצביע next לאיבר הבא ברשימה. כמו כן, אם השורה מתארתפקודת .word כל השדות כוללים את הערך NONO והשדה imm את מילת הזיכרון שיש לאחסן

// this struct will be used to save the memory lines

typedefstructMemoryLine {

// the opcode of the command

char opcode[6];

// the registers

char rd[6];

char rs[6];

char rt[6];

// immediate value

char imm[50];

// position of character in line

int pos;

// memory will also be a linked list to support infinite length programs

structMemoryLine\* next;

}MemoryLine;

הגדרה המבנה

גם למבנה זה ישנן פונקציות נלוות:

// creates new memory line with no "next"

MemoryLine\* create\_line(charopcode[6], charrd[6], charrs[6], charrt[6], charimm[50], intpos) {

// allocate memory for the label and create a pointer to it

MemoryLine\* new\_line = (MemoryLine\*)malloc(sizeof(MemoryLine));

// if allocation successful. insert data to label

if (new\_line != NULL) {

// use strcpy to insert the strings

strcpy(new\_line->opcode, opcode);

strcpy(new\_line->rd, rd);

strcpy(new\_line->rs, rs);

strcpy(new\_line->rt, rt);

strcpy(new\_line->imm, imm);

new\_line->pos = pos;

// no next defined

new\_line->next = NULL;

}

return new\_line;

}

ה"constructor" של המבנה, מקבל את ערכי כל השדות, מקצה זיכרון ואם הוא מצליח מציב את הערכים שהוא קיבל כך שהאיבר הבא ברשימה לא מוגדר

// adds line to memory structure. this line will be added to the end to let the writing run it like an array

MemoryLine\* add\_line(MemoryLine\* head, charopcode[6], charrd[6], charrs[6], charrt[6], charimm[50], intpos)

{

// the last line as for now

MemoryLine\* tail;

// create a line

MemoryLine\* new\_line = create\_line(opcode, rd, rs, rt, imm, pos);

// if the new line is null. do nothing

if (new\_line == NULL)

return NULL;

// and return the current if no head supplied

if (head == NULL)

return new\_line;

// get the "tail" to the end of the list

tail = head;

while (tail->next != NULL)

tail = tail->next;

// add the new line

tail->next = new\_line;

// return updated memory

Return head;

}

פונקציה שמוסיפה שורת זיכרון נוספת לרשימה. היא משתמשת במצביע tail על מנת להוסיף את השורה בסוף הרשימה. דבר שחשבנו שיקל על התוכנה בהמשך.

// destroy the memory line list and free the memory the assembler used

void destroy\_memLine(MemoryLine\* head) {

*// temp - a pointer to a line we are going to destroy after updating head*

MemoryLine\* temp;

*// all the way to the end*

**while** (head != NULL)

{

*// temp gets the current node while head advances*

temp = head;

head = head->next;

*// we destroy temp and free the memory*

free(temp);

}

}

דומה מאוד ל-destroy של הלייבלים. פונקציה זו לוקחת רשימה של שורות זיכרון, ומשחררת אותה מהזיכרון(של האסמבלר). המשתנה temp עוזר בשחרור כך שתמיד יהיה לנו מצביע לראש הרשימה הנוכחי

// get the memory line at position. can return null if does not exist

MemoryLine\* getAtPos(MemoryLine\* head, intpos) {

// go until you find

while (head != NULL&&head->pos != pos)

head = head->next;\_ returnhead;

}

פונקציה פשוטה מאוד שמקבלת מספר של שורת זיכרון ומחזירה מצביע לאיבר הרשימה שכולל את המידע עליה

מבנה עזר - זיכרון –Memory:

שדה זה כולל רשימה headשל מידע על שורות הזיכרון. ומספר last המכיל את מספר השורה האחרונה בזיכרון. מספר lastזה יהיה שווה למספר שורות הקוד שאינן הערה או לייבל במקרה שאין פקודות .word במקרה ופקודות כאלו קיימות הוא יתאר את הכתובת של האחרונה או הכתובת שורת הקוד האחרונה, מה שיותר גדול

// Memory struct and related functions. it is used so the second iteration can return two values.

typedefstructMemory

{

// head of memory line list

MemoryLine\* head;

// position of last

int last;

}Memory;

למבנה זה רק פונקציה שתי פונקציות. שכן השימוש בפקודות שרלוונטיות אליו מועט. פונקציה היוצרת זכרון ומחזירה אותו ופונקציה שמשחררת את הזיכרון של המבנה על ידי שחרור הרשימה

// destroys the memory struct after use

void destroy\_mem(Memory\* mem) {

// destroy the memory's line list

destroy\_memLine(mem->head);

// free the memory object's own memory

free(mem);

}

זוהי הפונקציה המשמידה את הזכרון על ידי שחרור הרשימה

*//create memory structure*

Memory\* create\_mem(MemoryLine\* head, int pos1)

{

Memory\* mem = (Memory\*)malloc(**sizeof**(Memory));

mem->head = head;

mem->last = pos1;

**return**(mem); *//Return number of lines*

}

וזוהי הפונקציה היוצרת זיכרון דינאמי ומחזירה אותו

איטרציות על הקבצים:

האסמבלר מבצע את עבודתו באיטרציות על הקבצים. שתי איטרציות על קובץ הקלט. אחת בה הוא משיג מידע על הלייבלים, ואחת בה הוא משיג מידע על שורות הזיכרון. אחרי שתי איטרציות אלו הוא מחליף את שמות הלייבלים היכן שצריך(בעיקר הוראות קפיצה) במספרי כתובות הזיכרון שלהם ורושם את תמונת הזיכרון בקובץ הפלט

איטרציה ראשונה – יצירת רשימת הלייבלים:

איטרציה זו עוברת על הקוד(אותו היא מקבלת במצביע asembl לקובץ הקלט). בודקת אם שורת קוד כוללת לייבל על ידי זיהוי התו ":" ואם כן מאחסנת מידע על הלייבל ברשימת הלייבלים

// the code of the first iteration. goes trough the file row by row and looks for labels, then adds them to the label list

label\* createLabelList(FILE \*asembl) {

הגדרות משתנים אשר נעשה בהם שימוש. הערת הקוד מעל מסבירה על תפקידו של כל משתנה.

*// rowIndex-the code row's index. where the PC will go after reading the label*

*// k is the char index for label name read, j is the index in the label name string we are building, option is determining if it's a label only line or a label + command line*

*// counter will be the line number in the new hexadecimal code. it will go up when a line that gets translated is found*

int rowIndex = 0, k, j, option, counter = 0;

*// line the current line being read, tav1 is the first char and i used to check for remarks, // tav - current char when reading label name*

*// label\_line will contain the name of the label once iteration is complete, dots are used to say "this is a label"*

char line[MAX\_LINE], tav1 ,tav, lable\_line[50], dots[50];

label\* head = NULL; *// the label list's head*

אחרי שסיימנו עם ההגדרות. נשתמש בלולאה while על מנת לקרוא את קובץ הקלט ולאתר לייבלים. הלולאה מסתיימת כאשר מתרחש אירוע feof של המצביע לקובץ

// go all the way trough the file

while (!feof(asembl)) {

דבר ראשון – קריאת השורה

// read a command from the assembler file

fgets(line, MAX\_LINE, asembl);

// reset option

option = 0;

איתור סוג השורה. במקרה של שורה ריקה או פקודת .word או הערה בתו הראשון עוברים לשורה הבאה

if (strcmp(line, "\n") == 0) //If line is blank, continue

continue;

tav1 = line[0];

if (tav1 == '#') //If line is Remark, continue

continue;

if (strstr(line, ".word") != NULL) //If line is .word, continue

continue;

אך אם אותר התו נקודותיים האסמבלר מבין שהוא הגיע לשורת לייבל

if (strstr(line, dots) != NULL) //If dots are found, this is a label

{

אך קודם הוא עורך בדיקה האם הנקודותיים נמצאות בתוך הערה מעבר לתו # ואז זו אזעקת שווא ועוברים לשורה הבאה

if (strstr(line, "#") != NULL) // however, ":" can be in a remark. so check for that as well, if so go to another line

if ((strstr(line, dots)) > (strstr(line, "#")))

continue;

קוראים את שם הלייבל. הקריאה עוברת על הקוד תו תו עד שהיא מגיעה לתו ":" ומתעתיקה את שם לייבל למחרוזת lable\_line

k = 0; j = 0; //Read the label name, first reset indexes

do {

// get current char

tav = line[k];

if (tav != ':') {

if (tav != '\t' && tav != ' ') *// don't read tabs and spaces*

{

// grab the read char and put it in name string

lable\_line[j] = tav;

// increment name string index

j++;

}

// increment reading index

k++;

}

} while (tav != ':');

בסוף. הלייבל צריך להיגמר בתו null כמו כל מחרוזת בסי לכן נוסיף אותו.

// label name is null terminated

lable\_line[j] = '\0';

חלק הקוד הבא מקדם את האינדקס בשורת קוד האסמבלי k על מנת לדעת אם שורת הזיכרון עליה הלייבל מצביע מתפרשת בקוד על שתי שורות קוד(לייבל ואחריו פקודה) או על שורה אחת שכוללת גם לייבל וגם פקודה על מנת שנוכל לקדם את אינדקס שורת הזיכרון counter בהתאם

k++; // Check if the line is lable line only by seeing if there are only spaces and tabs till the end

while ((line[k] == ' ') || (line[k] == '\t'))

k++;

// option is 1 on label only line, otherwise 0

if ((line[k] == '\n') || (line[k] == '#'))

option = 1;

אחרי הבדיקה. אנו מוסיפים את הלייבל לרשימת הלייבלים

// finally we add the label to label list

head = add(head, lable\_line, counter);

ובמקרה של לייבל בלבד. מחזירים את אינדקס כתובת הזיכרון אחד אחורה על מנת לא לקדם אותו בשורת הזיכרון

if (option == 1) { // Only label line - add label and decrement counter

counter = counter - 1;

}

}

לאחר ההצבה. נקדם את האינדקס k למיקום בו מתחילה הפקודה הבאה בתור הכנה אליה

k = 0; // Check if the current line is space line using k - most commands in fib.asm and our files start with a tab or a space

if ((line[k] == '\t') || (line[k] == ' '))

k++;

if (line[k] == '\n')

continue;

קידום אינדקס שורת הזיכרון

// increment hexa file line counter

counter++;

}

בסוף, נחזיר את רשימת הלייבלים

// return the list

return head;

}

איטרציה שניה – רשימת שורות קובץ הזיכרון:

פונקציות עזר לאיטרציה השנייה:

תפקיד הפונקציה specialworld היא לטפל בפקודות .word כלומר לשמור בזכרון.

הפונקציה שומרת את ערך המיקום בזכרון ואת הערך השמור על ידי מעבר תו אחרי תו.

Head- רשימת מיקומי השורות

Line- השורה אותה אנו קוראים עכשיו

Pose1- אינדקס מיקום סוף השורה

K – התו אותו אנו קוראים

MemoryLine \*specialword(MemoryLine\* head, char line[MAX\_LINE], int \*pos1, int \*k) {

char wordP[15], wordN[15]; *// wordP - address, wordN – data*

קודם בוצעה פעולת דילוג על רווחים.

int j = 0; *// index for string we copy to*

\*k = 0; *// reset k index*

**while** (line[\*k] != ' ') \*k=\*k+1;*// go past all the spaces*

\*k = \*k + 1;

j = 0; *//Copy Address. first reset j then copy char by char until the next space*

בחלק זה נבצע העתקה של מיקום הזכרון אל המערך המייצג את מיקום הזכרון

**while** (line[\*k] != ' ') {

wordP[j] = line[\*k];

j = j + 1; \*k = \*k + 1;

}

באותו אופן בדיוק מבוצעת העתקה של הערך שצריך לשמור למערך המייצג אותו.

wordP[j] = '\0'; \*k = \*k + 1; *// terminate string with null and increment end to next char*

j = 0; *//Copy Data. using the same way.*

**while** (line[\*k] != ' ')

{

**if** (line[\*k] == '\n') **break**;*// but detect an end of line string because after the data there can be a line end*

wordN[j] = line[\*k];

j = j + 1; \*k = \*k + 1;

}

wordN[j] = '\0';

בחלק זה מתבצעת קביעה האם הערך נתון כמספר דצימלי או הקסאדצימלי וטיפול לכל מקרה.

int pos = 0; *// pos - current line address, can be hexadecimal or decimal*

**if** (wordP[0] == '0') { *// change Address int. the if block considers an hexadecimal input*

**if** (wordP[1] == 'x' || wordP[1] == 'X') pos = strtol(wordP, NULL, 16);

}

**else** pos = atoi(wordP); *// and the else blocks considers a decimal input*

.word בתור אינדקציה להמשך שמדובר בפעולת NONOבחלק זה נבצע את השמירה זכרון עצמו, נתשמש ב

*// now. we will save the command in the memory list. NONO will be used as an indicator when writing to turn the command into a .word*

char nono[5] = "NONO"; *// a string used to copy nono to required places. fifth char is null*

strcpy(nono, "NONO"); strcpy(nono, "NONO"); strcpy(nono, "NONO"); strcpy(nono, "NONO");

head = add\_line(head, nono, nono, nono, nono, wordN, pos); *// save line to line list. wordN - the immediate value, is used as the data*

**if** (pos > \*pos1) \*pos1 = pos; *// update the location of the end of the memory*

**return** head;

לבסוף נחזיר את רשימת השורות העדכנית.

}

פונקציות readorder קוראת את הפקודה שוב על ידי מעבר תו אחרי תו , ושומרת אותה במערך עזר option.

void readorder(char line[MAX\_LINE], char \*option, int \*k){

char tav; *// current read char at index k*

int j = 0; *// index of copied char*

נבצע את הקריאה עד שנזהה כי הגענו לשימוש ברגיסטר

**do** { *// reading opcode should continue till dollar of first register*

tav = line[\*k]; *// read current*

**if** (tav != '$') *// if it's not dollar*

{

**if** (tav != '\t') *// or whitespace*

**if** (tav != ' ')

{

option[j] = tav; *// copy*

j = j + 1;

}

\*k = \*k + 1;

}

} **while** (tav != '$');

option[j] = '\0'; *// null terminate the opcode*

וכמובן בסוף נוסיף תו סיום כמו כל מחרוזת בשפת סי

הפונקציה הבאה תקדם אותנו עד לרגיסטר, פונקציה פשוטה המבצעת פעולה שחוזרת על עצמה.

}

void readdollar(char line[MAX\_LINE],int \*k){

**while** (line[\*k] != '$') \*k = \*k + 1; *// simply make your way to the dollar then stop*

}

פונקציה זו עוברת על הרגיסטר ושומרת את ערכו במערך. היא מקבלת את סוגו ובכך יודעת מה הוא, בסוף מחזירה את המערך.

*// reads register value*

*// line - current line being read*

*// rdst - register name. named so it can be used for rd, rs or rt*

void readrdst(char line[MAX\_LINE], char \*rdst, int \*k){

int j = 0; *// Read rd*

**while** (line[\*k] != ',')

{

**if** (line[\*k] != ' ' && line[\*k] != '\t') { *// read if not a whitespace*

rdst[j] = line[\*k];

j=j+1;

}

\*k=\*k+1;

}

rdst[j] = '\0';*// null terminate*

**return** rdst; *// return the string*

}

שומרת ומחזירה אותו במערךimmהפונקציה הבאה עוברת על ערך ה

void readimmd(char line[MAX\_LINE], char \*imm, int \*k){

*// go to immediate*

**while** ((line[\*k] == ' ') || (line[\*k] == '\t') || (line[\*k] == ',')) {

\*k = \*k + 1;

}

int j = 0; *// index of char being copied in immediate string*

**while** (line[\*k] != ' ')

{

**if** (line[\*k] != ' ' && line[\*k] != '\t') {

**if** ((line[\*k] == '\t') || (line[\*k] == '#') || (line[\*k] == '\n')) **break**;

imm[j] = line[\*k];

j=j+1;

}

\*k=\*k+1;

}

imm[j] = '\0';

**return** imm;

}

הפונקציה- readLine מבצעת את קריאת השורה לזכרון, היא משתמשת בפונקציות הקודומות ושומרת את התוצאות ברשימת שורות הזכרון(Head) ומחזירה אותה לאחר מכן.

*// reads line of memory and adds to memory line list "head*

*// line - what we read*

*// pos1 - number of last line of memin*

*// i - current instruction line index. might be more than pos1. is processed as we go so that's why a pointer*

*// head - the memory line list we add to*

*// k - index of char being read. is processed as we go so that's why a pointer*

MemoryLine \*readLine(char \*line, int \*pos1, int \*i, MemoryLine \*head, int \*k) {

char option[6], rd[6], rs[6], rt[6], imm[50]; *// the line's properties*

readorder(line, option, k); *// read the opcode*

readrdst(line, rd, k); *// Read rd*

readdollar(line, k); *// wait for dollar sign*

readrdst(line, rs, k); *// Read rs*

readdollar(line, k); *// wait for dollar sign*

readrdst(line, rt, k); *// Read rt*

readimmd(line, imm, k); *//handle immediate*

head = add\_line(head, option, rd, rs, rt, imm, \*i); \*i = \*i + 1; *// save line to line list*

**if** (\*i > \*pos1) \*pos1 = \*i; *//Update last line position*

**return** head;

}

כעת הגענו אל האיטרציה השנייה ממש, איטרציה זו גם כן עוברת על קוד האסמבלי, אך היא בודקת עבור כל שורה בקובץ איך היא משפיעה על קובץ הזיכרון. ומכניסה את המידע לרשימת שורות הזיכרון. המצביע file הינו מצביע לקובץ הקלט

Memory\* SecondRun(FILE\* file) {

*// k - the index of the current char being read, i - the current position in the file*

*// pos1 - the last line of the memory file*

int k = 0, i = 0, pos1 = 0;

*// char - line will house the current line. tav1 will save the first character of the line and option, rd, rs, rt and are the command's values*

*// dots - used to detect labels. because something might be past them*

char line[MAX\_LINE], tav1, \* dots = ":";

MemoryLine\* head = NULL; *// the Memory list's head. it will contain info about each memory line in the end*

**while** ((fgets(line, MAX\_LINE, file)) != NULL) { *// the loop reads the file line by line. and upon reaching null it stops as that's where the file ends*

**if** (strcmp(line, "**\n**") == 0) **continue**; *// in case of a Blank line, go*

tav1 = line[0]; *// get first line*

**if** (tav1 == '#') **continue**; *// in case of a Remark line, go*

עד חלק זה מתבצעת הגדרת משתנים ודילוג על שורות ריקות או בעלות הערות בלבד

char wo[6] = ".word"; *// a string for comparison*

int isword = 0; *// booleand for .world detection*

ערכים הבודקים האם זוהי פקודה מיוחדת של.word במקרה וזוהי אכן פקודה מיוחדת השורות הבאות יטפלו בה

**if** (strstr(line, wo) != NULL) { *// in case of the special .word order*

head = specialword(head, line, &pos1, &k);

isword = 1;

}

עבור פקודות רגילות עם לייבל נבדוק האם הנקודותיים הם הערה ולא לייבל, במקרה וזוהי אכן שורות לייבל נבדוק גם האם זוהי רק שורת לייבל או גם הוראה

**else** **if** (strstr(line, dots) != NULL){ *//in case of regular order and label*

**if** (strstr(line, "#") != NULL){ *//now we check if the dots is remark and not a label*

**if** ((strstr(line, dots)) >= (strstr(line, "#"))) { *//label line- check if the line include only label or order*

**goto** mark; *// start reading*

}

} *// the following code section will work if the there is mark # and it appears after dots or if there isn't mark #*

k = 0;

**while** (line[k] != ':') k++;

k++;

**if** (line[k] == '\n') **continue**;

**else**

**while** ((line[k] == ' ') || (line[k] == '\t')) k++;

**if** (line[k] == '\n') **continue**;

**if** (line[k] == '#') **continue**;

**if** (i > pos1) pos1 = i; *//Update last line position*

}

**else***// Order line only*

{

k = 0;

**while** ((line[k] == ' ') || (line[k] == '\t')) k++; *// roll to end of spaces*

**if** (line[k] == '#') **continue**;

**if** (line[k] == '\n') **continue**;

}

חלק זה מבצע את העתקת השורה במקרה וזוהי לא .word command

**if** (!isword) { *// copy line in all not .world scenarios*

mark:

head = readLine(line, &pos1, &i, head, &k);

}

}

ולבסוף ניצור את הזכרון של הקובץ.

**return** create\_mem(head, pos1); *// create memory structure and return it to*

*main function*

}

הפונקציות הבאות עוזרות להדפיס לקובץ הסופי את הדרוש:

הפונקציה pringrdrsrt עוזרת לנו להדפיס את שמות הרגיסטרים, בפועל ירשם מספר המייצג את שם הרגיסטר.

*// rdst - register name to print*

*// memin - current file pointer*

*// num - will be flipped to zero if it's a .world*

void printrdrsrt(char \*rdst, FILE \*memin, int \*num)

{ *// basically a big if block that checks the name of the register and converts it to a number*

**if** (strcmp(rdst, "$zero") == 0)

fprintf(memin, "0");

**else** **if** (strcmp(rdst, "$imm") == 0)

fprintf(memin, "1");

**else** **if** (strcmp(rdst, "$v0") == 0)

fprintf(memin, "2");

**else** **if** (strcmp(rdst, "$a0") == 0)

fprintf(memin, "3");

**else** **if** (strcmp(rdst, "$a1") == 0)

fprintf(memin, "4");

**else** **if** (strcmp(rdst, "$t0") == 0)

fprintf(memin, "5");

**else** **if** (strcmp(rdst, "$t1") == 0)

fprintf(memin, "6");

**else** **if** (strcmp(rdst, "$t2") == 0)

fprintf(memin, "7");

**else** **if** (strcmp(rdst, "$t3") == 0)

fprintf(memin, "8");

**else** **if** (strcmp(rdst, "$s0") == 0)

fprintf(memin, "9");

**else** **if** (strcmp(rdst, "$s1") == 0)

fprintf(memin, "A");

**else** **if** (strcmp(rdst, "$s2") == 0)

fprintf(memin, "B");

**else** **if** (strcmp(rdst, "$gp") == 0)

fprintf(memin, "C");

**else** **if** (strcmp(rdst, "$sp") == 0)

fprintf(memin, "D");

**else** **if** (strcmp(rdst, "$fp") == 0)

fprintf(memin, "E");

**else** **if** (strcmp(rdst, "$ra") == 0)

fprintf(memin, "F");

**else** **if** (strcmp(rdst, "NONO") == 0)

\*num = 0;

**else**

fprintf(memin, "0");

}

הפונקציה הבאה מדפיסה את הפקודה (שוב על ידי מספר המייצג את סוג הפעולה) ולבסוף מחזירה ערך בוליאני שאחראי להודיע לנו האם בוצעה הדפסה.

*// opc - opcode string*

*// memin - output file pointer*

int printopcode(char \*opc, FILE \*memin)

{

**if** (strcmp(opc, "add") == 0) {

fprintf(memin, "00"); **return** 1;}

**else** **if** (strcmp(opc, "sub") == 0) {

fprintf(memin, "01"); **return** 1;}

**else** **if** (strcmp(opc, "and") == 0) {

fprintf(memin, "02"); **return** 1;}

**else** **if** (strcmp(opc, "or") == 0) {

fprintf(memin, "03"); **return** 1;}

**else** **if** (strcmp(opc, "sll") == 0) {

fprintf(memin, "04"); **return** 1;}

**else** **if** (strcmp(opc, "sra") == 0) {

fprintf(memin, "05"); **return** 1;}

**else** **if** (strcmp(opc, "srl") == 0) {

fprintf(memin, "06"); **return** 1;}

**else** **if** (strcmp(opc, "beq") == 0) {

fprintf(memin, "07"); **return** 1;}

**else** **if** (strcmp(opc, "bne") == 0) {

fprintf(memin, "08"); **return** 1;}

**else** **if** (strcmp(opc, "blt") == 0) {

fprintf(memin, "09"); **return** 1;}

**else** **if** (strcmp(opc, "bgt") == 0) {

fprintf(memin, "0A"); **return** 1;}

**else** **if** (strcmp(opc, "ble") == 0) {

fprintf(memin, "0B"); **return** 1;}

**else** **if** (strcmp(opc, "bge") == 0) {

fprintf(memin, "0C"); **return** 1;}

**else** **if** (strcmp(opc, "jal") == 0) {

fprintf(memin, "0D"); **return** 1;}

**else** **if** (strcmp(opc, "lw") == 0) {

fprintf(memin, "0E"); **return** 1;}

**else** **if** (strcmp(opc, "sw") == 0) {

fprintf(memin, "0F"); **return** 1;}

**else** **if** (strcmp(opc, "reti") == 0) {

fprintf(memin, "10"); **return** 1;}

**else** **if** (strcmp(opc, "in") == 0) {

fprintf(memin, "11"); **return** 1;}

**else** **if** (strcmp(opc, "out") == 0) {

fprintf(memin, "12"); **return** 1;}

**else** **if** (strcmp(opc, "halt") == 0) {

fprintf(memin, "13"); **return** 1;}

**else** *// on .word*

**return** 0;

}

הפונקציה PrintDataToFile מבצעת את ההדפסה לקובץ. היא משתמשת בהרבה מפונקציות העזר על מנת לבצע את ההדפסה. היא עוברת שורה שורה עד לאחרונה בקובץ. בכל איטרציה היא מדפיסה את הפקודה, אם ואין פקודה היא תדפיס 00 ואילו אם לא התבצעה כתיבה(מצביע על .word command) תטפל גם במקרה זה. לאחר מכן היא עוברת להדפיס את הרגיסטרים, עבור כל אחד מהם היא בודקת האם ישנו אם לא היא תדפיס 0 ואם כן תדפיס את הרגיסטר. לבסוף היא מטפלת במקרה של פקודה מיוחדת של .word

*//gets memory head and output file indicator. prints the memory into file*

void PrintDataToFile(Memory\* mem, FILE \*memin)

{

*// i - memory index, num - word for .word*

int i = 0, num = 0; int flag=0;

**while** (mem->head != NULL && i <= mem->last) {

MemoryLine \*currentLine = getAtPos(mem->head, i); *// get the current line's data once. this will reduce the code's execution time. allowing it to build apps much more quickly*

*// Printing Opcode. if data for the ith row does not exist print a zero*

**if** (currentLine == NULL) fprintf(memin, "00");*// if no opcode print 2 zeros*

**else** flag=printopcode(currentLine->opcode, memin); *// print the opcode and return if it was printed*

**if** (!flag && currentLine != NULL) {*// if there is no opcode. this block of code is used to get the word for the .word command*

**if** ((strcmp(currentLine->opcode, "NONO") == 0)) {

**if** ((currentLine->imm[0] == '0') && ((currentLine->imm[1] == 'x') || (currentLine->imm[1] == 'X')))

num = strtol(currentLine->imm, NULL, 16);

**else** *//Imiddiate is decimal*

num = atoi(currentLine->imm);

fprintf(memin, "%08X", num); *//Print immidiate in hex*

}

}

**else** **if** (!flag) *// if there is nothing print a zero*

fprintf(memin, "00");

**if** (currentLine == NULL) fprintf(memin, "0"); *// Printing Rd*

**else** printrdrsrt(currentLine->rd, memin,&num);

**if** (currentLine == NULL) fprintf(memin, "0"); *// Printing Rs*

**else** printrdrsrt(currentLine->rs, memin,&num);

**if** (currentLine == NULL) fprintf(memin, "0"); *// Printing Rt*

**else** printrdrsrt(currentLine->rt, memin,&num);

*// a check wheter to print the immediate and skip .word lines*

**if** (currentLine == NULL) fprintf(memin, "%03X", 0 & 0xfff); *// on a null line. print zero to immediate*

**else** **if**(strcmp(currentLine->opcode, "NONO") != 0) *// now print if satisfied*

{

**if** ((currentLine->imm[0] == '0') && ((currentLine->imm[1] == 'x') || (currentLine->imm[1] == 'X'))) num = strtol(currentLine->imm, NULL, 16); *//Check if immidiate in hex*

**else** num = atoi(currentLine->imm);

fprintf(memin, "%03X", num & 0xfff); *//Print immidiate in hex. the & 0xfff is supposed to shorte negative numbers to 3 hexadecimal digits or 12 bits*

}

**if** (i != mem->last) fprintf(memin, "**\n**"); *//Print \n except the last line*

i++; *// go to next line*

}

}

הפונקציה LableChange דואגת לבצע החלפה בין שם הלייבל למיקמו המספרי בקובץ. בנוסף היא מבצעת החלפה בין שם רגיסטר האפס ($zero) ל0.

*// a label switch function that runs between the second run and the write. changes label names in the memory structure to their locations taken*

*// from the label structure*

*// this function also changes immediate to zero if the register name $zero was recorded in the immediate field*

void LableChange(MemoryLine\* head, label\* lb)

{

char temp[50];

*// the current memory line*

MemoryLine \*current = head;

**while** (current != NULL) {

*// find if there is a label on the immediate and if it exists*

int loc = find(lb,current->imm);

*// if found*

**if** (loc != -1) {

\_itoa(loc, temp, 10); *// Changes int to string and puts in temp*

strcpy(current->imm, temp); *// Copy label location number to immidiate*

}

**if** (strcmp(current->imm, "$zero") == 0) *// If immidiate is &zero*

{

strcpy(current->imm, "0"); *// Changes immidiate to "0"*

}

current = current->next;

}

}

ולבסוף יש לנו את החלק המרכזי המבצע את הפעולה כולה על ידי קריאה לפונקציות העזר:

הוא פותח את הקובץ. אחר כך הוא מבצע את הקריאה הראשונה, כלומר יצירת רשימת הלייבלים.

לאחר מכן הוא מבצע את האיטרציה השנייה שהיא בניית הזכרון והחלפת שמות הלייבלים לכתובות מתאימות בזכרון, וכמובן סוגרת את הקובץ לאחר מכן. לבסוף התוכנית פותחת את הקובץ אליו היא כותבת, מבצעת את הכתיבה וסוגרת את הקובץ. בפעולות האחרונות מתבצעות שחרור הזכרון הדינאמי שיצרנו.

*// part 3 - the main function*

*// the main takes two arguments, the input file and the output file. indexes start with 1 because argv[0] is the program itself*

int main(int argc, char\* argv[]) {

*// open the input file. doing so in the main function will allow us to have infinite length file names*

*// why i call it "asembl"? because of what it is*

FILE \*asembl = fopen(argv[1], "r");

*// leave if null file is supplied*

**if** (asembl == NULL) {

exit(1);

}

*// the first iteration, locate the labels and write thier locations to the linked list*

label\* labels = createLabelList(asembl);

*// close the file from the first iteration*

fclose(asembl);

*// and reopen it for the second*

asembl = fopen(argv[1], "r");

*// another null check in case something happend*

**if** (asembl == NULL) {

exit(1);

}

*// start the second iteration*

Memory \*memory = SecondRun(asembl);

fclose(asembl);

LableChange(memory->head, labels); *// Change labels from words to numbers*

*// Write Data to file*

FILE\* memin = fopen(argv[2], "w");

**if** (memin == NULL)

exit(1);

PrintDataToFile(memory, memin);

fclose(memin);

*// End of file writing*

*// free the memory taken by the label list and memory structure*

destroy(labels);

destroy\_mem(memory);

סוגרים אותו כדי שתוכנות אחרות יוכלו להיכנס אליו(כגון הסימולטור)

fclose(memin);

// End of file writing

ולבסוף, מנקים את כל הזיכרון של המבנים(סוג של garbage collection ידני)

// free the memory taken by the label list and memory structure

destroy(labels);

destroy\_mem(memory);