

MIPS-Programmierung in der WebSPIM-Umgebung (0.3)

C. Reichenbach, <mailto:reichenbach@cs.uni-frankfurt.de>

22. April 2014

1 Einführung

WebSPIM ist ein Web-basierter MIPS32-Simulator, dessen MIPS-Funktionalität auf dem Simulator „SPIM“¹, entwickelt von James Larus (U. Wisconsin) basiert. WebSPIM erweitert SPIM wie folgt:

- Web-basierte Nutzerschnittstelle
- Übungsaufgaben-Management
- Übungsspezifische Speicherfunktion
- Kustomisierungen des simulierten Systems je nach Übungsaufgabe

Dieses Dokument ist eine Kurzbeschreibung zur Verwendung von SPIM. Abschnitt 2 erläutert WebSPIM-Spezifika, während Abschnitt 3 eine Beschreibung der von SPIM verwendeten Assemblersprache ist.

2 WebSPIM

Das WebSPIM-System besteht aus drei wesentlichen Komponenten:

- Der Aufgabenauswahlkomponente
- Der Programmmeditor
- Der Simulator

2.1 Aufgabenauswahl

Dies ist die Einstiegsseite des Systems. Die Aufgabe zeigt die möglichen Interaktionen mit dem System an, je nach Konfiguration. Dies kann z.B. die freie Interaktion mit dem Simulator sein, oder auch spezifische Aufgaben mit einem festgesetzten spätesten Abgabetermin.

2.2 Programmmeditor

Der Editor (Abbildung 2) besteht aus einem Textfeld mit Assemblercode (linker Hand), einem Block mit der Aufgabenstellung (rechts oben) und (bei Übungen) einen Einsendebereich im rechten unteren Bereich.

Der Editor selbst hat nur grundlegendste Programmmanipulationsfähigkeiten. Es kann (je nach persönlicher Präferenz) einfacher sein, in einem anderen Editor zu arbeiten und den erzeugten Assemblercode dann in den WebSPIM-Editor zu kopieren.

Der Einsendebereich ist nur bei Übungen verfügbar, deren letzter Abgabetermin noch nicht abgelaufen ist. Bei bereits abgelaufenen Übungen wird hier nur die zuletzt abgegebene Übung angezeigt.

Wenn die Übung noch eingereicht werden kann, besteht dieser Eingabebereich aus zwei weiteren Eingabefeldern. Eines dient dazu, eventuelle Namen von mitarbeitenden Studierenden anzugeben. Das zweite, größere Feld kann für verschiedenartige Anmerkungen genutzt werden. Rechts davon ist die aktuellst eingesehene Fassung von Programm, Anmerkungen, und Mitarbeitern. Dieser Eintrag kann durch Drücken des „Einsenden“-Knopfes überschrieben werden, solange der letzte Abgabetermin noch nicht überschritten ist.

Der „Programm Laden“-Knopf links unterhalb dem Programmmeditor lädt das Programm in den Simulator, ohne es einzusenden.

¹ www.cs.wisc.edu/~larus/spim.html

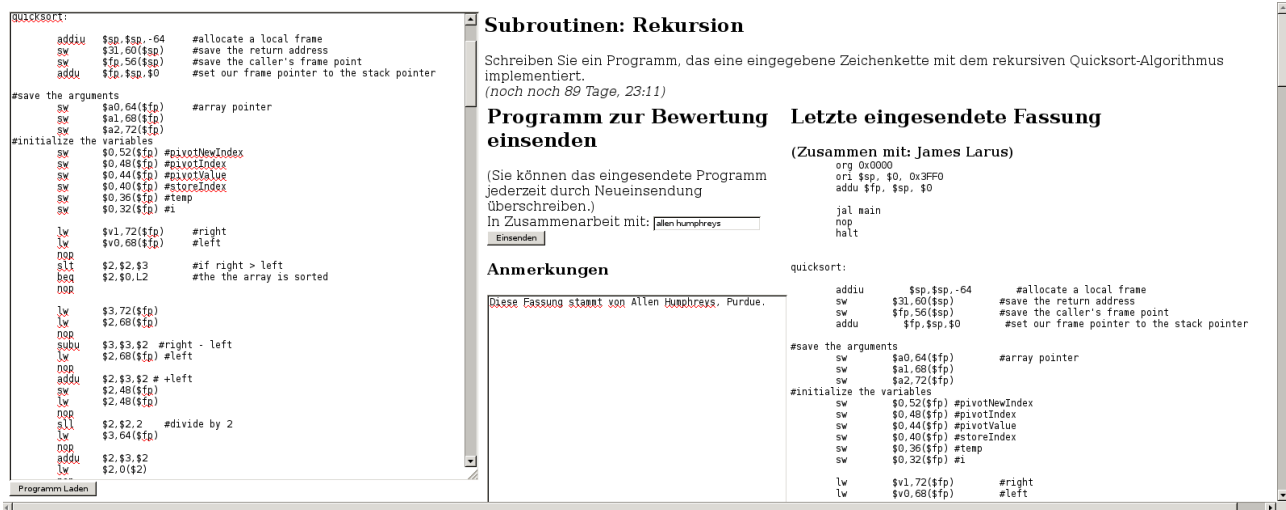


Abbildung 1: Der WebSPIM-Editor.

2.3 Simulator

Der Simulator besteht aus sechs Komponenten:

- Im linken Drittel ist eine Speicheranzeige.
- Im mittleren Drittel ist eine Anzeige der Assemblerbefehle.
- Rechts oben sind Links, die zurück zum Hauptmenü oder zurück zum Programmeditor springen. Wenn das laufende Programm vom Editor aus nicht eingeseordnet wurde, wird es beim Sprung in das Hauptmenü verworfen, beim Sprung in den Programmeditor aber wieder hergestellt.
- Darunter ist eine Anzeige der aktuellen Registerinhalte.
- Darunter ist eine kleine Leiste mit Ausführknöpfen.
- Ganz rechts unten ist die Konsolenausgabe des Systems. Hier werden Ausgaben des Prozessors gesammelt.

2.3.1 Die Speicheranzeige

Ganz links in eckigen Klammern sind Adressen wie z.B. [00400000] angegeben. Diese Adressen geben (hexadezimal notiert) die Speicheradresse des ersten der sechzehn folgenden Bytes (ebenfalls hexadezimal notiert) an. Eine Zeile der Form

[00400000] 01 02 03 04 05 06 07 08 00 00 00 00 0A 0B 0C 0D

besagt also, daß das Byte an Adresse 00400000₁₆ im Speicher den Wert 01₁₆ hat, das Wort an Adresse 00400001₁₆ den Wert 02₁₆ usw.

2.3.2 Die Assembleranzeige

Im mittleren Drittel sind alle identifizierten Assemblerbefehle aufgeführt, zusammen mit ihrer Speicheradresse und ihrem Maschinencode. Ein Doppelpfeil (⇒) markiert die Instruktion, auf die der aktuelle Befehlszähler (PC) verweist.

2.3.3 Ausführknöpfe

Die beiden Ausführknöpfe können zur Einzelschrittausführung oder zur Mehrschrittausführung verwendet werden; bei letzterer muß die Anzahl der gewünschten Schritte angegeben werden. Der Einzelausführungsknopf kann durch Drücken der s-Taste direkt ausgelöst werden.

Die Simulationsumgebung beobachtet Änderungen von Speicher- und Registerinhalten und markiert diese in der Anzeige durch hellgrüne Unterlegung des geänderten Wertes.

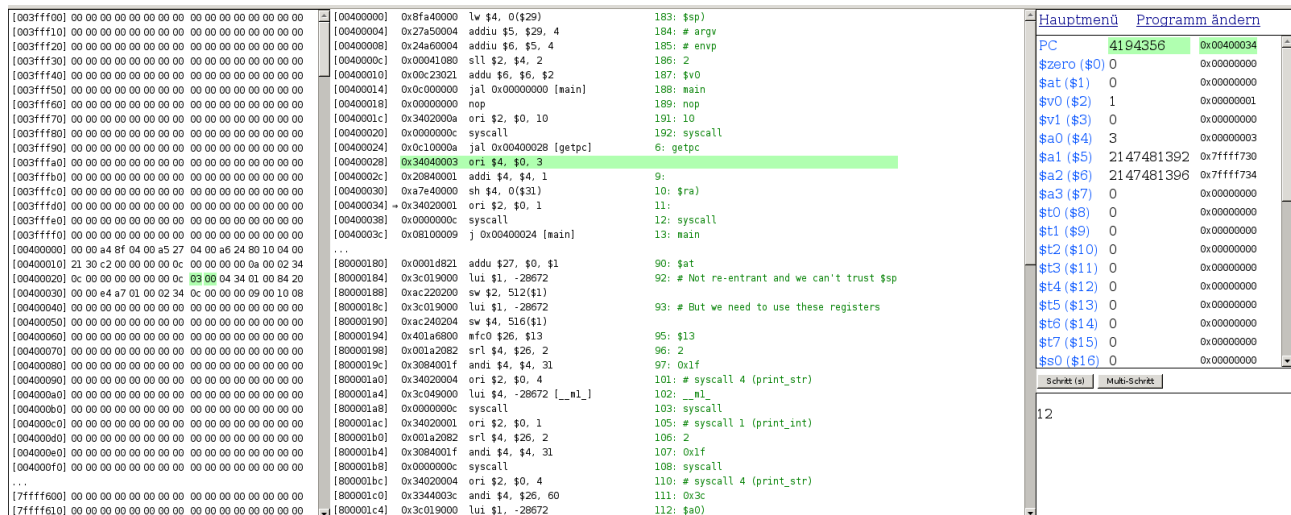


Abbildung 2: Das WebSPIM-Interface zu SPIM.

3 MIPS-Programmierung mit SPIM

Der SPIM-Assembler ist Teil des SPIM-Systems und somit auch Teil von WebSPIM. Seine Aufgabe ist es, Assemblerbefehle in Maschinensprache zu übersetzen, um diese ausführen zu können.

Maschinenspracheprogramme bestehen aus vier Bausteinen:

- Direktiven, die Daten angeben oder Eigenschaften des Programms erläutern
- Marken, als Ziele für Sprung-/Verzweigungsoperationen (Sprungmarken) oder auch für Speicheroperationen (Speichermarken)
- Assemblerbefehle
- Kommentare

Hier ist ein kleines SPIM-Programm, das diese Konzepte illustriert:

```
.text          # Markiert den folgenden Teil als Assembler
.globl main   # Markiert Sprungmarke "main" als global sichtbar

main:
    li $a0, 2
    li $v0, 1
    syscall
    li $v0, 10
    syscall
```

Die erste Zeile beinhaltet eine Direktive `.text`. Direktive beginnen immer mit einem vorangestellten Punkt. Der Rest der Zeile ist Kommentar, eingeleitet von einem Lattenzaun-Zeichen (`#`): Dieses Zeichen markiert immer das erste Zeichen eines Kommentars, so daß der Rest der Zeile automatisch vom Assembler ignoriert wird.

Zeile Drei besteht aus der Sprungmarke `main`. Sprungmarken werden durch einen Doppelpunkt abgeschlossen. Sie werden vom Assembler verwendet, um die Adresse eines Objektes zu bestimmen. Sie können sowohl relativ (z.B. bei `bgz`, `lw`) als auch absolut (z.B. bei `j`, `jal`) verwendet werden; der Assembler erkennt dies automatisch.

Die letzten fünf Zeilen schließlich sind Assemblerbefehle.

Wir beschränken uns in dieser Beschreibung auf die grundsätzlichen Fähigkeiten des SPIM-Assemblers, wie sie zum Lösen typischer Programmieraufgaben nötig sind.

3.1 Direktiven

3.1.1 Sektionsmarkierungen

Diese Markierungen geben an, in welchen Teil des Speichers dieser Teil des Assemblerprogrammes geladen werden soll.

.data : Diese Direktive instruiert den Assembler, die folgenden Teile in das Datensegment zu laden. Typischerweise werden hier globale Variablen reserviert, z.B. per **.byte**, **ascii**, oder **.word**-Direktive.

.text : Diese Direktive instruiert den Assembler, die folgenden Teile in das Textsegment zu laden, also in den zur Programmausführung designierten Speicher.

3.1.2 Sichtbarkeit von Marken

Sprung- und Speichermarken sind normalerweise nur innerhalb des eigenen Moduls sichtbar. In WebSPIM ist das geladene Programm immer ein eigenständiges Modul, somit sind dessen Marken vom Rest des Systems streng getrennt. Dies kann jedoch über einige Deklarationen gesteuert werden.

.globl : Diese Direktive wird zusammen mit einer Marke verwendet, z.B.

```
.globl main
```

Sie macht diese Marke global sichtbar, erlaubt also anderen Modulen, sie anzuspringen oder von ihr zu laden. Die Sprungmarke `textmain` muß im Hauptmodul immer global markiert sein, damit das Betriebssystem weiß, wohin im Programm es springen soll, um das Programm zu starten.

.comm, **.lcomm** : Diese Direktiven allozieren Speicher (üblicherweise im **.data**-Bereich), geben dem Speicher eine Marke und setzen die Sichtbarkeit. **.comm** setzt globale Sichtbarkeit, während **.lcomm** lokale Sichtbarkeit setzt. Die Direktiven haben zwei Parameter, durch ein Komma getrennt: der Name der Marke, und der reservierte Platz. Zum Beispiel:

```
.lcomm puffer, 40
```

alloziert 40 Bytes an einer Marke **puffer**.

.set : Diese Direktive ändert das Verhalten des Assemblers bezüglich Assemblercode, der auf das **\$at**-Register zugreift.

```
.set at
```

verbietet dem Programmierer, dieses Register direkt zu verwenden, aber erlaubt Pseudobefehle, die das Register verwenden.

```
.set noat
```

verbietet dem Programmierer, Pseudobefehle zu verwenden, die **\$at** verwenden, aber erlaubt direkten Zugriff auf dieses Register.

3.1.3 Platzallozierungen

Allozierungen finden normalerweise im **.data**-Segment statt. Diese Operationen stellen eine gewisse Menge an Speicher im aktuellen Segment zur Verfügung. Einige von ihnen initialisieren den Speicher auch (d.h., sie setzen den Speicher auf vordefinierte Werte).

.space : alloziert die angegebene Menge an Speicher, initialisiert diesen aber nicht.

```
.space 8
```

alloziert zum Beispiel 8 Bytes.

.byte : alloziert ein Byte für jedes Byte, das dem **.byte** folgt (durch Komma getrennt). Beispiel:

```
.byte 12, 0x0a
```

alloziert zwei Bytes mit den Werten 12 (**0xc**) und 10 (**0xa**).

.half, **.word** : Diese beiden Direktiven funktionieren wie **.byte**, allerdings für Halbwoorte und Worte.

.ascii, **.asciiz** : Diese Direktiven allozieren Zeichenketten. Beide legen ASCII-Zeichen ab; der Unterschied ist, daß **asciiz** hinter diese Zeichen noch ein Null-Byte legt, zum Beispiel:

```
.asciiz "Hallo, Welt!"
```

3.2 Assemblerbefehle

Wir unterscheiden zwischen Hardwarebefehlen und Pseudobefehlen. Pseudobefehle markieren wir mit einem unten angestellten "p", z.B. **li_p**.

3.2.1 Arithmetische und logische Befehle

abs_p \$z, \$x: $z := |x|$ (d.h., es wird der Betrag berechnet.)

add \$z, \$x, \$y: $z := x + y$. Ausnahmebehandlung bei Überlauf.

addu \$z, \$x, \$y: $z := x + y$. Überlauf wird abgeschnitten.

addi \$z, \$x, v: $z := x + v$. Ausnahmebehandlung bei Überlauf. **v** ist vorzeichenbehaftet.

addiu \$z, \$x, v: $z := x + v$. Überlauf wird abgeschnitten. **v** ist nicht vorzeichenbehaftet.

and \$z, \$x, \$y: $z := x \& y$. Bitweise Und-Verknüpfung der Bits.

andi \$z, \$x, v: $z := x \& v$. Bitweise Und-Verknüpfung der Bits.

clo \$z, \$x: Zählt die führenden Einsen in der Bit-Repräsentierung des Inhaltes von **\$x**.

clz \$z, \$x: Zählt die führenden Nullen in der Bit-Repräsentierung des Inhaltes von **\$x**.

div \$x, \$y: $LO := \lfloor \frac{x}{y} \rfloor$, $HI := x \bmod y$. Ausnahmebehandlung bei Überlauf. Registerinhalte werden vorzeichenbehaftet interpretiert. Der Divisionsrest in **HI** ist nur für positive Operanden definiert, bei negativen Operanden gibt die Befehlssatzarchitektur keine Garantien zum Wert von **HI**.

divu \$x, \$y: $LO := \lfloor \frac{x}{y} \rfloor$, $HI := x \bmod y$. Überlauf wird abgeschnitten. Registerinhalte werden ohne Vorzeichen interpretiert.

div_p \$z, \$x, \$y: $z := \lfloor \frac{x}{y} \rfloor$. Ausnahmebehandlung bei Überlauf. Vorzeichenbehaftet.

divu_p \$z, \$x, \$y: $z := \lfloor \frac{x}{y} \rfloor$. Überlauf wird abgeschnitten. Nicht vorzeichenbehaftet.

mult \$x, \$y: $HI:LO := x \times y$. Registerinhalte werden vorzeichenbehaftet interpretiert.

multu \$x, \$y: $HI:LO := x \times y$. Registerinhalte werden ohne Vorzeichen interpretiert.

mul_p \$z, \$x, \$y: $z := x \times y$. Überlauf wird abgeschnitten. Registerinhalte werden vorzeichenbehaftet interpretiert.

mulo_p \$z, \$x, \$y: $z := x \times y$. Ausnahmebehandlung bei Überlauf. Registerinhalte werden vorzeichenbehaftet interpretiert.

mulou_p \$z, \$x, \$y: $z := x \times y$. Ausnahmebehandlung bei Überlauf. Registerinhalte werden ohne Vorzeichen interpretiert.

neg_p \$z, \$x: $z = -x$. Arithmetische Negation. Ausnahmebehandlung bei Überlauf.

negu_p \$z, \$x: $z = -x$. Arithmetische Negation. Überlauf wird ignoriert.

nor \$z, \$x, \$y: $z := \neg(x | y)$. Bitweise Nicht-Oder-Verknüpfung der Bits.

or \$z, \$x, \$y: $z := x | y$. Bitweise Oder-Verknüpfung der Bits.

ori \$z, \$x, v: $z := x | v$. Bitweise Oder-Verknüpfung der Bits.

rem_p \$z, \$x, \$y: $z := x \bmod y$. Ausnahmebehandlung bei Überlauf. Vorzeichenbehaftet.

remu_p \$z, \$x, \$y: $z := x \bmod y$. Überlauf wird abgeschnitten. Nicht vorzeichenbehaftet.

sll \$z, \$x, v: $z := x \ll v$. Verschiebt Bits nach links und füllt mit Nullen auf.

sllv \$z, \$x, \$y: $z := x \ll y$. Verschiebt Bits nach links und füllt mit Nullen auf.

sra \$z, \$x, v: $z := x \gg^- v$. Verschiebt Bits nach rechts und füllt mit Vorzeichenbit auf.

srav \$z, \$x, \$y: $z := x \gg^- y$. Verschiebt Bits nach rechts und füllt mit Vorzeichenbit auf.

srl \$z, \$x, v: $z := x \gg v$. Verschiebt Bits nach rechts und füllt mit Nullen auf.

srlv \$z, \$x, \$y: $z := x \gg y$. Verschiebt Bits nach rechts und füllt mit Nullen auf.

rol_p \$z, \$x, v: Verschiebt Bits nach links und füllt rechts mit den hinausgeschobenen Bits auf.

ror_p \$z, \$x, \$y: Verschiebt Bits nach rechts und füllt links mit den hinausgeschobenen Bits auf.

sub \$z, \$x, \$y: $z := x - y$. Ausnahmebehandlung bei Überlauf.

subu \$z, \$x, \$y: $z := x - y$. Überlauf wird abgeschnitten.

xor \$z, \$x, \$y: $z := x \wedge y$. Bitweise Exklusiv-Oder-Verknüpfung der Bits.

xori \$z, \$x, v: $z := x \wedge v$. Bitweise Exklusiv-Oder-Verknüpfung der Bits.

3.2.2 Ladeoperationen

move_p \$z, \$x: $z := x$

li_p \$z, v: Lade 32 Bit-Wert v in Register z

la_p \$z, ziel: Lade Adresse der Sprungmarke $ziel$ in Register z

lui \$z, v: Lade obere 16 Bit von z mit v , setze untere 16 Bit auf 0.

mfhi \$z: $z := HI$

mflo \$z: $z := LO$

mthi \$z: $HI := z$

mtlo \$z: $LO := z$

3.2.3 Vergleichsbefehle

slt \$z, \$x, \$y: Setzt \$z auf 1 gdw $x < y$ (mit Vorzeichen), sonst 0.

sltu \$z, \$x, \$y: Setzt \$z auf 1 gdw $x < y$ (ohne Vorzeichen), sonst 0.

slti \$z, \$x, v: Setzt \$z auf 1 gdw $x < v$ (mit Vorzeichen), sonst 0.

sltiu \$z, \$x, v: Setzt \$z auf 1 gdw $x < v$ (ohne Vorzeichen), sonst 0.

seq_p \$z, \$x, \$y: Setzt \$z auf 1 gdw $x = y$ (mit Vorzeichen), sonst 0.

sge_p \$z, \$x, \$y: Setzt \$z auf 1 gdw $x \geq y$ (mit Vorzeichen), sonst 0.

sgeu_p \$z, \$x, \$y: Setzt \$z auf 1 gdw $x \geq y$ (ohne Vorzeichen), sonst 0.

sgt_p \$z, \$x, \$y: Setzt \$z auf 1 gdw $x > y$ (mit Vorzeichen), sonst 0.

sgtu_p \$z, \$x, \$y: Setzt \$z auf 1 gdw $x > y$ (ohne Vorzeichen), sonst 0.

sle_p \$z, \$x, \$y: Setzt \$z auf 1 gdw $x \leq y$ (mit Vorzeichen), sonst 0.

sleu_p \$z, \$x, \$y: Setzt \$z auf 1 gdw $x \leq y$ (ohne Vorzeichen), sonst 0.

sne_p \$z, \$x, \$y: Setzt \$z auf 1 gdw $x \neq y$ (mit Vorzeichen), sonst 0.

3.2.4 Sprung- und Verzweigungsbefehle

b ziel: $PC := \text{ziel}$

bnez \$x, ziel: $PC := \text{ziel}$ falls $x \neq 0$

beqz \$x, ziel: $PC := \text{ziel}$ falls $x = 0$

beq \$x, \$y, ziel: $PC := \text{ziel}$ falls $x = y$

bne \$x, \$y, ziel: $PC := \text{ziel}$ falls $x \neq y$

bgez \$x, ziel: $PC := \text{ziel}$ falls $x \geq 0$ (mit Vorzeichen)

bgezal \$x, ziel: $\$ra := PC + 4; PC := \text{ziel}$ falls $x \geq 0$ (mit Vorzeichen)

bgtz \$x, ziel: $PC := \text{ziel}$ falls $x > 0$ (mit Vorzeichen)

blez \$x, ziel: $PC := \text{ziel}$ falls $x \leq 0$ (mit Vorzeichen)

bltz \$x, ziel: $PC := \text{ziel}$ falls $x < 0$ (mit Vorzeichen)

bltzal \$x, ziel: $\$ra := PC + 4; PC := \text{ziel}$ falls $x < 0$ (mit Vorzeichen)

bge_p \$x, \$y, ziel: $PC := \text{ziel}$ falls $x \geq y$ (mit Vorzeichen)

bgeu_p \$x, \$y, ziel: PC := ziel falls $x \geq y$ (ohne Vorzeichen)

bgt_p \$x, \$y, ziel: PC := ziel falls $x > y$ (mit Vorzeichen)

bgtu_p \$x, \$y, ziel: PC := ziel falls $x > y$ (ohne Vorzeichen)

ble_p \$x, \$y, ziel: PC := ziel falls $x \leq y$ (mit Vorzeichen)

bleu_p \$x, \$y, ziel: PC := ziel falls $x \leq y$ (ohne Vorzeichen)

blt_p \$x, \$y, ziel: PC := ziel falls $x < y$ (mit Vorzeichen)

bltu_p \$x, \$y, ziel: PC := ziel falls $x < y$ (ohne Vorzeichen)

j ziel: PC := ziel

jal ziel: \$ra := PC + 4; PC := ziel

jr \$x: PC := \$x

jalr \$x: \$ra := PC + 4; PC := \$x

3.2.5 Speicheroperationen (Laden und Schreiben)

Lade- und Schreibbefehle für den Speicher greifen jeweils auf Adresse $v + x$ zu. Sie addieren also den Inhalt von Register x mit dem Direktoperanden v und legen dadurch die Speicheradresse fest, auf die sie schreiben bzw. von der sie lesen. Solche Zugriffe müssen entsprechend den Bedürfnissen des Befehls im Speicher ausgerichtet sein, sonst wird eine Ausnahmebehandlung im Betriebssystemkern angestoßen.

Ladebefehle:

Befehl	Name	Ausrichtung	B	Vorzeichenerweiterung
lb \$z, v(\$x)	load byte	8	1	ja
lbu \$z, v(\$x)	load byte unsigned	8	1	nein
lh \$z, v(\$x)	load halfword	16	2	ja
lhu \$z, v(\$x)	load halfword unsigned	16	2	nein
lw \$z, v(\$x)	load word	32	4	—

Schreibbefehle:

Befehl	Name	Ausrichtung	Bytes
sb \$z, v(\$x)	store byte	8	1
sh \$z, v(\$x)	store halfword	16	2
sw \$z, v(\$x)	store word	32	4

3.2.6 Systemaufrufe

syscall : Springe in die Unterbrechungsbehandlungsroutine des Betriebssystems und bitte um einen Systemdienst. Der genaue Systemdienst wird durch eine Kennzahl in Register $v0$ ausgewählt. Die verfügbaren Systemdienste variieren je nach Konfiguration von WebSPIM. Hier sind die häufigst unterstützten Dienste:

\$v0	Parameter	Beschreibung	Wirkung
1	\$a0	Zahl ausgeben	Druckt die Zahl in \$a0 auf der Ausgabekonsolle aus.
4	\$a0	Zeichenkette ausgeben	Druckt die null-terminierte ASCII -Zeichenkette, die im Speicher ab Adresse \$a0 abgelegt ist, auf der Ausgabekonsolle aus.
5		Zahl einlesen	Liest eine Zahl ein und schreibt diese in Register \$v0 .
8	\$a0, \$a1	Zeichenkette einlesen	Liest eine null-terminierte ASCII -Zeichenkette ein. Die Zeichenkette wird ab Adresse \$a0 in den Speicher geschrieben. Wenn mehr als (\$a1-1) Zeichen eingegeben wurden, wird die Zeichenkette gekürzt: maximal \$a1 Zeichen werden geschrieben.
10		System anhalten	Bbeauftragt das Betriebssystem, den Prozessor anzuhalten.

3.3 Register

Register	Name	Bedeutung	Erhalten?
	PC	Programmzähler	—
\$0	\$zero	Konstante 0	—
\$1	\$at	Assembler-Hilfsregister	<i>nein</i>
\$2–\$3	\$v0–\$v1	Rückgabewerte	<i>nein</i>
\$4–\$7	\$a0–\$a3	Argumente (Parameter)	<i>nein</i>
\$8–\$15	\$t0–\$t7	Temporäre Register	<i>nein</i>
\$16–\$23	\$s0–\$s7	Gesicherte Register	<i>ja</i>
\$24–\$25	\$t8–\$t9	Temporäre Register	<i>nein</i>
\$26–\$27	\$k0–\$k1	Kernelregister	<i>nein</i>
\$28	\$gp	Globaler Zeiger	<i>ja</i>
\$29	\$sp	Stapelzeiger	<i>ja</i>
\$30	\$fp	Rahmenzeiger	<i>ja</i>
\$31	\$ra	Rücksprungadresse	—
	HI	Mult.-Einheit, obere 32 Bit	<i>nein</i>
	LO	Mult.-Einheit, untere 32 Bit	<i>nein</i>