

new/usr/src/cmd/xargs/xargs.c

1

```
*****
24711 Mon Mar 31 18:53:08 2014
new/usr/src/cmd/xargs/xargs.c
4703 would like xargs support for -P
*****

1 /*
2  * CDDL HEADER START
3  *
4  * The contents of this file are subject to the terms of the
5  * Common Development and Distribution License (the "License").
6  * You may not use this file except in compliance with the License.
7  *
8  * You can obtain a copy of the license at usr/src/OPENSOLARIS.LICENSE
9  * or http://www.opensolaris.org/os/licensing.
10 * See the License for the specific language governing permissions
11 * and limitations under the License.
12 *
13 * When distributing Covered Code, include this CDDL HEADER in each
14 * file and include the License file at usr/src/OPENSOLARIS.LICENSE.
15 * If applicable, add the following below this CDDL HEADER, with the
16 * fields enclosed by brackets "[]" replaced with your own identifying
17 * information: Portions Copyright [yyyy] [name of copyright owner]
18 *
19 * CDDL HEADER END
20 */
21 /*
22 * Copyright 2012 DEY Storage Systems, Inc. All rights reserved.
23 *
24 * Portions of this file developed by DEY Storage Systems, Inc. are licensed
25 * under the terms of the Common Development and Distribution License (CDDL)
26 * version 1.0 only. The use of subsequent versions of the License are
27 * is specifically prohibited unless those terms are not in conflict with
28 * version 1.0 of the License. You can find this license on-line at
29 * http://www.illumos.org/license/CDDL
30 */
31 /*
32 * Copyright 2008 Sun Microsystems, Inc. All rights reserved.
33 * Use is subject to license terms.
34 */

36 /*      Copyright (c) 1984, 1986, 1987, 1988, 1989 AT&T */
37 /*      All Rights Reserved      */

40 #include <stdio.h>
41 #include <sys/types.h>
42 #include <sys/wait.h>
43 #include <unistd.h>
44 #include <fcntl.h>
45 #include <string.h>
46 #include <stdarg.h>
47 #include <stdlib.h>
48 #include <limits.h>
49 #include <wchar.h>
50 #include <locale.h>
51 #include <langinfo.h>
52 #include <stropts.h>
53 #include <poll.h>
54 #include <errno.h>
55 #include <stdarg.h>
56 #include "getresponse.h"

58 #define HEAD    0
59 #define TAIL    1
60 #define FALSE  0
61 #define TRUE   1
```

new/usr/src/cmd/xargs/xargs.c

2

```
62 #define MAXSBUF 255
63 #define MAXIBUF 512
64 #define MAXINSERTS 5
65 #define BUFSIZE LINE_MAX
66 #define MAXARGS 255
67 #define INSPAT_STR    "{}"      /* default replstr string for -[Ii] */
68 #define FORK_RETRY    5

70 #define QBUF_STARTLEN 255 /* start size of growable string buffer */
71 #define QBUF_INC 100     /* how much to grow a growable string by */

73 /* We use these macros to help make formatting look "consistent" */
74 #define MSG(s)      ermsg(gettext(s "\n"))
75 #define MSG2(s, a)  ermsg(gettext(s "\n"), a)
76 #define PERR(s)     perror(gettext("xargs: " s))

78 /* Some common error messages */

80 #define LIST2LONG      "Argument list too long"
81 #define ARG2LONG       "A single argument was greater than %d bytes"
82 #define MALLOCFAIL     "Memory allocation failure"
83 #define CORRUPTFILE    "Corrupt input file"
84 #define WAITFAIL       "Wait failure"
85 #define CHILDSIG       "Child killed with signal %d"
86 #define CHILDFAIL      "Command could not continue processing data"
87 #define FORKFAIL       "Could not fork child"
88 #define EXECFAIL       "Could not exec command"
89 #define MISSQUOTE      "Missing quote"
90 #define BADESCAPE       "Incomplete escape"
91 #define IBUFOVERFLOW    "Insert buffer overflow"
92 #define NOCHILDSLOT     "No free child slot available"
93 #endif /* ! codereview */

95 #define _(x)      gettext(x)

97 static wctype_t blank;
98 static char      *arglist[MAXARGS+1];
99 static char      argbuf[BUFSIZE * 2 + 1];
100 static char      lastarg[BUFSIZE + 1];
101 static char      **ARGV = arglist;
102 static char      *LEOF = "_";
103 static char      *INSPAT = INSPAT_STR;
104 static char      ins_buf[MAXIBUF];
105 static char      *p_ibuf;

107 static struct inserts {
108     char      **p_ARGV;      /* where to put newarg ptr in arg list */
109     char      *p_skel;      /* ptr to arg template */
110 } saveargv[MAXINSERTS];

112 static int        PROMPT = -1;
113 static int        BUFLIM = BUFSIZE;
114 static int        MAXPROCS = 1;
115 #endif /* ! codereview */
116 static int        N_ARGS = 0;
117 static int        N_args = 0;
118 static int        N_lines = 0;
119 static int        DASHX = FALSE;
120 static int        MORE = TRUE;
121 static int        PER_LINE = FALSE;
122 static int        ERR = FALSE;
123 static int        OK = TRUE;
124 static int        LEGAL = FALSE;
125 static int        TRACE = FALSE;
126 static int        INSERT = FALSE;
127 static int        ZERO = FALSE;
```

```

128 static int      linesize = 0;
129 static int      ibufsize = 0;
130 static int      exitstat = 0; /* our exit status */
131 static int      mac; /* modified argc, after parsing */
132 static char     **mav; /* modified argv, after parsing */
133 static int      n_inserts; /* # of insertions. */
134 static pid_t    *procs; /* pids of children */
135 static int      n_procs; /* # of child processes. */
136 #endif /* ! codereview */

138 /* our usage message: */
139 #define USAGEMSG "Usage: xargs: [-t] [-p] [-0] [-e[EOFSTR]] [-E EOFSTR] \"\
140 \"[-I replstr] [-i[replstr]] [-L #] [-l[#]] [-n #] [-x]] [-P maxprocs] [-s\
92 \"[-I replstr] [-i[replstr]] [-L #] [-l[#]] [-n #] [-x]] [-s size] \"\
141 \"[cmd [args ...]]\\n\"

143 static int      echoargs();
144 static wint_t   getchrc(char *, size_t *);
145 static void     lcall(char *sub, char **subargs);
97 static int     lcall(char *sub, char **subargs);
146 static void     addibuf(struct inserts *p);
147 static void     errmsg(char *messages, ...);
148 static char     *addarg(char *arg);
149 static void     store_str(char **, char *, size_t);
150 static char     *getarg(char *);
151 static char     *insert(char *pattern, char *subst);
152 static void     usage();
153 static void     parseargs();
154 static void     procs_malloc(void);
155 static int      procs_find(pid_t child);
156 static void     procs_store(pid_t child);
157 static int      procs_delete(pid_t child);
158 static pid_t    procs_waitpid(int blocking, int *stat_loc);
159 static void     procs_wait(int blocking);
160 #endif /* ! codereview */

162 int
163 main(int argc, char **argv)
164 {
165     int      j;
166     struct inserts *psave;
167     int c;
168     int      initsize;
169     char     *cmdname, **initlist;
170     char     *arg;
171     char     *next;

173     /* initialization */
174     blank = wctype("blank");
175     n_inserts = 0;
176     psave = saveargv;
177     (void) setlocale(LC_ALL, "");
178     #if !defined(TEXT_DOMAIN) /* Should be defined by cc -D */
179     #define TEXT_DOMAIN "SYS_TEST" /* Use this only if it weren't */
180     #endif
181     (void) textdomain(TEXT_DOMAIN);
182     if (init_yes() < 0) {
183         errmsg(_(ERR_MSG_INIT_YES), strerror(errno));
184         exit(1);
185     }

187     parseargs(argc, argv);

189     /* handling all of xargs arguments: */
190     while ((c = getopt(mac, mav, "0tpe:E:I:l:n:P:s:x")) != EOF) {
191         while ((c = getopt(mac, mav, "0tpe:E:I:l:n:P:s:x")) != EOF) {

```

```

191     switch (c) {
192     case '0':
193         ZERO = TRUE;
194         break;

196     case 't': /* -t: turn trace mode on */
197         TRACE = TRUE;
198         break;

200     case 'p': /* -p: turn on prompt mode. */
201         if ((PROMPT = open("/dev/tty", O_RDONLY)) == -1) {
202             PERR("can't read from tty for -p");
203         } else {
204             TRACE = TRUE;
205         }
206         break;

208     case 'e':
209         /*
210          * -e[EOFSTR]: set/disable end-of-file.
211          * N.B. that an argument *isn't* required here; but
212          * parseargs forced an argument if not was given. The
213          * forced argument is the default...
214          */
215         LEOF = optarg; /* can be empty */
216         break;

218     case 'E':
219         /*
220          * -E EOFSTR: change end-of-file string.
221          * EOFSTR *is* required here, but can be empty:
222          */
223         LEOF = optarg;
224         break;

226     case 'I':
227         /* -I replstr: Insert mode. replstr *is* required. */
228         INSERT = PER_LINE = LEGAL = TRUE;
229         N_ARGS = 0;
230         INSPAT = optarg;
231         if ((*optarg == '\0') {
232             errmsg_(_("Option requires an argument: -%c\\n"),
233                 c);
234         }
235         break;

237     case 'i':
238         /*
239          * -i [replstr]: insert mode, with *optional* replstr.
240          * N.B. that an argument *isn't* required here; if
241          * it's not given, then the string INSPAT_STR will
242          * be assumed.
243          *
244          * Since getopt(3C) doesn't handle the case of an
245          * optional variable argument at all, we have to
246          * parse this by hand:
247          */

249         INSERT = PER_LINE = LEGAL = TRUE;
250         N_ARGS = 0;
251         if ((optarg != NULL) && (*optarg != '\0')) {
252             INSPAT = optarg;
253         } else {
254             /*
255              * here, there is no next argument. so
256              * we reset INSPAT to the INSPAT_STR.

```

```

257         * we *have* to do this, as -i/I may have
258         * been given previously, and XCU4 requires
259         * that only "the last one specified takes
260         * effect".
261         */
262         INSPAT = INSPAT_STR;
263     }
264     break;

266 case 'L':
267     /*
268      * -L number: # of times cmd is executed
269      * number *is* required here:
270      */
271     PER_LINE = TRUE;
272     N_ARGS = 0;
273     INSERT = FALSE;
274     if ((PER_LINE = atoi(optarg)) <= 0) {
275         errmsg(_("#lines must be positive int: %s\n"),
276             optarg);
277     }
278     break;

280 case 'l':
281     /*
282      * -l [number]: # of times cmd is executed
283      * N.B. that an argument *isn't* required here; if
284      * it's not given, then l is assumed.
285      *
286      * parseargs handles the optional arg processing.
287      */

289     PER_LINE = LEGAL = TRUE; /* initialization */
290     N_ARGS = 0;
291     INSERT = FALSE;

293     if ((optarg != NULL) && (*optarg != '\0')) {
294         if ((PER_LINE = atoi(optarg)) <= 0)
295             PER_LINE = 1;
296     }
297     break;

299 case 'n': /* -n number: # stdin args */
300     /*
301      * -n number: # stdin args.
302      * number *is* required here:
303      */
304     if ((N_ARGS = atoi(optarg)) <= 0) {
305         errmsg(_("#args must be positive int: %s\n"),
306             optarg);
307     } else {
308         LEGAL = DASHX || N_ARGS == 1;
309         INSERT = PER_LINE = FALSE;
310     }
311     break;

313 case 'P': /* -P maxprocs: # of child processes */
314     MAXPROCS = atoi(optarg);
315     if (MAXPROCS <= 0) {
316         errmsg(_("#maxprocs must be positive int: %s\n"),
317             optarg);
318     }
319     break;

321 #endif /* ! codereview */
322 case 's': /* -s size: set max size of each arg list */

```

```

323         BUFLIM = atoi(optarg);
324         if (BUFLIM > BUFSIZE || BUFLIM <= 0) {
325             errmsg(_("0 < max-cmd-line-size <= %d: %s\n"),
326                 BUFSIZE, optarg);
327         }
328         break;

330 case 'x': /* -x: terminate if args > size limit */
331     DASHX = LEGAL = TRUE;
332     break;

334 default:
335     /*
336      * bad argument. complain and get ready to die.
337      */
338     usage();
339     exit(2);
340     break;
341 }
342 }

344 /*
345  * if anything called errmsg(), something screwed up, so
346  * we exit early.
347  */
348 if (OK == FALSE) {
349     usage();
350     exit(2);
351 }

353 /*
354  * we're finished handling xargs's options, so now pick up
355  * the command name (if any), and it's options.
356  */

359 mac -= optind; /* dec arg count by what we've processed */
360 mav += optind; /* inc to current mav */

362 (void) procs_malloc();

364 #endif /* ! codereview */
365 if (mac <= 0) { /* if there're no more args to process, */
366     cmdname = "/usr/bin/echo"; /* our default command */
367     *ARGV++ = addarg(cmdname); /* use the default cmd. */
368 } else { /* otherwise keep parsing rest of the string. */
369     /*
370      * note that we can't use getopt(3C), and *must* parse
371      * this by hand, as we don't know apriori what options the
372      * command will take.
373      */
374     cmdname = *mav; /* get the command name */

377 /* pick up the remaining args from the command line: */
378 while ((OK == TRUE) && (mac-- > 0)) {
379     /*
380      * while we haven't crapped out, and there's
381      * work to do:
382      */
383     if (INSERT && ! ERR) {
384         if (strstr(*mav, INSPAT) != NULL) {
385             if (++n_inserts > MAXINSERTS) {
386                 errmsg(_("too many args "
387                     "with %s\n"), INSPAT);
388                 ERR = TRUE;

```

```

389         }
390         psave->p_ARGV = ARGV;
391         (psave++)->p_skel = *mav;
392     }
393     }
394     *ARGV++ = addarg(*mav++);
395 }
396
397 /* pick up args from standard input */
398
399 initlist = ARGV;
400 initsize = linesize;
401 lastarg[0] = '\0';
402
403 while (OK) {
404     N_args = 0;
405     N_lines = 0;
406     ARGV = initlist;
407     linesize = initsize;
408     next = argbuf;
409
410     while (MORE || (lastarg[0] != '\0')) {
411         int l;
412
413         if (*lastarg != '\0') {
414             arg = strcpy(next, lastarg);
415             *lastarg = '\0';
416         } else if ((arg = getarg(next)) == NULL) {
417             break;
418         }
419
420         l = strlen(arg) + 1;
421         linesize += l;
422         next += l;
423
424         /* Inserts are handled specially later. */
425         if ((n_inserts == 0) && (linesize >= BUFLIM)) {
426             /*
427              * Legal indicates hard fail if the list is
428              * truncated due to size. So fail, or if we
429              * cannot create any list because it would be
430              * too big.
431              */
432             if (LEGAL || N_args == 0) {
433                 EMSG(LIST2LONG);
434                 (void) procs_wait(1);
435             }
436 #endif /* ! codereview */
437             exit(2);
438             /* NOTREACHED */
439         }
440
441         /*
442          * Otherwise just save argument for later.
443          */
444         (void) strcpy(lastarg, arg);
445         break;
446     }
447
448     *ARGV++ = arg;
449
450     N_args++;
451
452     if ((PER_LINE && N_lines >= PER_LINE) ||
453         (N_ARGS && (N_args) >= N_ARGS)) {
454         break;

```

```

455     }
456
457     if ((ARGV - arglist) == MAXARGS) {
458         break;
459     }
460 }
461
462 *ARGV = NULL;
463 if (N_args == 0) {
464     /* Reached the end with no more work. */
465     break;
466     exit(exitstat);
467 }
468
469 /* insert arg if requested */
470
471 if (!ERR && INSERT) {
472     p_ibuf = ins_buf;
473     ARGV--;
474     j = ibufsize = 0;
475     for (psave = savearg; ++j <= n_inserts; ++psave) {
476         addibuf(psave);
477         if (ERR)
478             break;
479     }
480 }
481 *ARGV = NULL;
482
483 if (n_inserts > 0) {
484     /*
485      * if we've done any insertions, re-calculate the
486      * linesize. bomb out if we've exceeded our length.
487      */
488     linesize = 0;
489     for (ARGV = arglist; *ARGV != NULL; ARGV++) {
490         linesize += strlen(*ARGV) + 1;
491     }
492     if (linesize >= BUFLIM) {
493         EMSG(LIST2LONG);
494         (void) procs_wait(1);
495     }
496 #endif /* ! codereview */
497     exit(2);
498     /* NOTREACHED */
499 }
500
501 /* exec command */
502
503 if (!ERR) {
504     if (!MORE &&
505         (PER_LINE && N_lines == 0 || N_ARGS && N_args == 0))
506         exit(exitstat);
507     OK = TRUE;
508     j = TRACE ? echoargs() : TRUE;
509     if (j) {
510         /*
511          * for xcu4, all invocations of cmdname must
512          * return 0, in order for us to return 0.
513          * so if we have a non-zero status here,
514          * quit immediately.
515          */
516         (void) lcall(cmdname, arglist);
517         exitstat |= lcall(cmdname, arglist);
518     }

```

```

519     }
520 }

522 (void) procs_wait(1);

524 #endif /* ! codereview */
525 if (OK)
526     return (exitstat);

528 /*
529  * if exitstat was set, to match XCU4 compliance,
530  * return that value, otherwise, return 1.
531  */
532 return (exitstat ? exitstat : 1);
533 }

535 static char *
536 addarg(char *arg)
537 {
538     linesize += (strlen(arg) + 1);
539     return (arg);
540 }

543 static void
544 store_str(char **buffer, char *str, size_t len)
545 {
546     (void) memcpy(*buffer, str, len);
547     (*buffer)[len] = '\0';
548     *buffer += len;
549 }

552 static char *
553 getarg(char *arg)
554 {
555     char *xarg = arg;
556     wchar_t c;
557     char mbc[MB_LEN_MAX];
558     size_t len;
559     int escape = 0;
560     int inquote = 0;

562     arg[0] = '\0';

564     while (MORE) {

566         len = 0;
567         c = getwchr(mbc, &len);

569         if (((arg - xarg) + len) > BUFLIM) {
570             MSG2(ARG2LONG, BUFLIM);
571             exit(2);
572             ERR = TRUE;
573             return (NULL);
574         }

576         switch (c) {
577             case '\n':
578                 if (ZERO) {
579                     store_str(&arg, mbc, len);
580                     continue;
581                 }
582                 /* FALLTHRU */

584             case '\0':

```

```

585         case WEOF: /* Note WEOF == EOF */

587             if (escape) {
588                 MSG(BADESCAPE);
589                 ERR = TRUE;
590                 return (NULL);
591             }
592             if (inquote) {
593                 MSG(MISSQUOTE);
594                 ERR = TRUE;
595                 return (NULL);
596             }

598             N_lines++;
599             break;

601         case '"':
602             if (ZERO || escape || (inquote == 1)) {
603                 /* treat it literally */
604                 escape = 0;
605                 store_str(&arg, mbc, len);

607             } else if (inquote == 2) {
608                 /* terminating double quote */
609                 inquote = 0;

611             } else {
612                 /* starting quoted string */
613                 inquote = 2;
614             }
615             continue;

617         case '\'':
618             if (ZERO || escape || (inquote == 2)) {
619                 /* treat it literally */
620                 escape = 0;
621                 store_str(&arg, mbc, len);

623             } else if (inquote == 1) {
624                 /* terminating single quote */
625                 inquote = 0;

627             } else {
628                 /* starting quoted string */
629                 inquote = 1;
630             }
631             continue;

633         case '\\':
634             /*
635              * Any unquoted character can be escaped by
636              * preceding it with a backslash.
637              */
638             if (ZERO || inquote || escape) {
639                 escape = 0;
640                 store_str(&arg, mbc, len);
641             } else {
642                 escape = 1;
643             }
644             continue;

646         default:
647             /* most times we will just want to store it */
648             if (inquote || escape || ZERO || !iswctype(c, blank)) {
649                 escape = 0;
650                 store_str(&arg, mbc, len);

```

```

651         continue;
652     }
653     /* unquoted blank */
654     break;
655 }
656
657 /*
658  * At this point we are processing a complete argument.
659  */
660 if (strcmp(xarg, LEOF) == 0 && *LEOF != '\0') {
661     MORE = FALSE;
662     return (NULL);
663 }
664 if (c == WEOF) {
665     MORE = FALSE;
666 }
667 if (xarg[0] == '\0')
668     continue;
669 break;
670 }
671
672 return (xarg[0] == '\0' ? NULL : xarg);
673 }
674
675 /*
676  * errmsg(): print out an error message, and indicate failure globally.
677  *
678  * Assumes that message has already been gettext()'d. It would be
679  * nice if we could just do the gettext() here, but we can't, since
680  * since xgettext(1M) wouldn't be able to pick up our error message.
681  */
682 /* PRINTFLIKE1 */
683 static void
684 errmsg(char *messages, ...)
685 {
686     va_list ap;
687
688     va_start(ap, messages);
689
690     (void) fprintf(stderr, "xargs: ");
691     (void) vfprintf(stderr, messages, ap);
692
693     va_end(ap);
694     OK = FALSE;
695 }
696
697 static int
698 echoargs()
699 {
700     char **anarg;
701     char **tanarg; /* tmp ptr */
702     int i;
703     char reply[LINE_MAX];
704
705     tanarg = anarg = arglist-1;
706
707     /*
708      * write out each argument, separated by a space. the tanarg
709      * nonsense is for xcu4 testsuite compliance - so that an
710      * extra space isn't echoed after the last argument.
711      */
712     while (++tanarg) { /* while there's an argument */
713         ++tanarg; /* follow anarg */
714         (void) write(2, *anarg, strlen(*anarg));
715
716         if (++tanarg) { /* if there's another argument: */

```

```

717         (void) write(2, " ", 1); /* add a space */
718         --tanarg; /* reset back to anarg */
719     }
720 }
721 if (PROMPT == -1) {
722     (void) write(2, "\n", 1);
723     return (TRUE);
724 }
725
726 (void) write(2, "?...", 4); /* ask the user for input */
727
728 for (i = 0; i < LINE_MAX && read(PROMPT, &reply[i], 1) > 0; i++) {
729     if (reply[i] == '\n') {
730         if (i == 0)
731             return (FALSE);
732         break;
733     }
734 }
735 reply[i] = 0;
736
737 /* flush remainder of line if necessary */
738 if (i == LINE_MAX) {
739     char bitbucket;
740
741     while ((read(PROMPT, &bitbucket, 1) > 0) && (bitbucket != '\n'))
742         ;
743 }
744
745 return (yes_check(reply));
746 }
747
748 static char *
749 insert(char *pattern, char *subst)
750 {
751     static char buffer[MAXSBUF+1];
752     int len, ipatlen;
753     char *pat;
754     char *bufend;
755     char *pbuf;
756
757     len = strlen(subst);
758     ipatlen = strlen(INSPT) - 1;
759     pat = pattern - 1;
760     pbuf = buffer;
761     bufend = &buffer[MAXSBUF];
762
763     while ((*++pat) {
764         if (strncmp(pat, INSPAT, ipatlen) == 0) {
765             if (pbuf + len >= bufend) {
766                 break;
767             } else {
768                 (void) strcpy(pbuf, subst);
769                 pat += ipatlen;
770                 pbuf += len;
771             }
772         } else {
773             *pbuf++ = *pat;
774             if (pbuf >= bufend)
775                 break;
776         }
777     }
778 }
779
780 if (!*pat) {
781     *pbuf = '\0';
782     return (buffer);

```

```

783     } else {
784         errmsg(gettext("Maximum argument size with insertion via %s's "
785             "exceeded\n"), INSPAT);
786         ERR = TRUE;
787         return (NULL);
788     }
789 }

792 static void
793 addibuf(struct inserts *p)
794 {
795     char    *newarg, *skel, *sub;
796     int      l;

798     skel = p->p_skel;
799     sub = *ARGV;
800     newarg = insert(skel, sub);
801     if (ERR)
802         return;

804     l = strlen(newarg) + 1;
805     if ((ibufsize += l) > MAXIBUF) {
806         EMSG(IBUFOVERFLOW);
807         ERR = TRUE;
808     }
809     (void) strcpy(p_ibuf, newarg);
810     *(p->p_ARGV) = p_ibuf;
811     p_ibuf += l;
812 }

815 /*
816 * getwchr():  get the next wide character.
817 * description:
818 *     we get the next character from stdin.  This returns WEOF if no
819 *     character is present.  If ZERO is set, it gets a single byte instead
820 *     a wide character.
821 */
822 static wint_t
823 getwchr(char *mbc, size_t *sz)
824 {
825     size_t    i;
826     int       c;
827     wchar_t   wch;

829     i = 0;
830     while (i < MB_CUR_MAX) {
832         if ((c = fgetc(stdin)) == EOF) {
834             if (i == 0) {
835                 /* TRUE EOF has been reached */
836                 return (WEOF);
837             }
839             /*
840              * We have some characters in our buffer still so it
841              * must be an invalid character right before EOF.
842              */
843             break;
844         }
845         mbc[i++] = (char)c;

847         /* If this succeeds then we are done */
848         if (ZERO) {

```

```

849         *sz = i;
850         return ((char)c);
851     }
852     if (mbtowc(&wch, mbc, i) != -1) {
853         *sz = i;
854         return ((wint_t)wch);
855     }
856 }

858 /*
859 * We have now encountered an illegal character sequence.
860 * There is nothing much we can do at this point but
861 * return an error.  If we attempt to recover we may in fact
862 * return garbage as arguments, from the customer's point
863 * of view.  After all what if they are feeding us a file
864 * generated in another locale?
865 */
866 errno = EILSEQ;
867 PERR(CORRUPTFILE);
868 exit(1);
869 /* NOTREACHED */
870 }

873 static void
874 static int
875 lcall(char *sub, char **subargs)
876 {
877     int      retry = 0;
878     pid_t    child;
879     int      retcode, retry = 0;
880     pid_t    iwait, child;

881     for (;;) {
882         switch (child = fork()) {
883             default:
884                 (void) procs_store(child);
885                 (void) procs_wait(0);
886                 return;
887                 while ((iwait = wait(&retcode)) != child &&
888                     iwait != (pid_t)-1)
889                     ;
890                 if (iwait == (pid_t)-1) {
891                     PERR(WAITFAIL);
892                     exit(122);
893                     /* NOTREACHED */
894                 }
895                 if (WIFSIGNALED(retcode)) {
896                     EMSG2(CHILDSIG, WTERMSIG(retcode));
897                     exit(125);
898                     /* NOTREACHED */
899                 }
900                 if ((WEXITSTATUS(retcode) & 0377) == 0377) {
901                     EMSG(CHILDFAIL);
902                     exit(124);
903                     /* NOTREACHED */
904                 }
905                 return (WEXITSTATUS(retcode));
906             case 0:
907                 (void) execvp(sub, subargs);
908                 PERR(EXECFAIL);
909                 if (errno == EACCES)
910                     exit(126);
911                 exit(127);
912                 /* NOTREACHED */
913             case -1:

```

```

893         if (errno != EAGAIN && retry++ < FORK_RETRY) {
894             PERR(FORKFAIL);
895             exit(123);
896         }
897         (void) sleep(1);
898     }
899 }
900 }

902 static void
903 procs_malloc(void)
904 {
905     int i;

907     procs = (pid_t *) (malloc(MAXPROCS * sizeof(pid_t)));
908     if (procs == NULL) {
909         PERR(MALLOCFAIL);
910         exit(1);
911     }

913     for (i = 0; i < MAXPROCS; i++) {
914         procs[i] = (pid_t)(0);
915     }
916 }

918 static int
919 procs_find(pid_t child)
920 {
921     int i;

923     for (i = 0; i < MAXPROCS; i++) {
924         if (procs[i] == child) {
925             return (i);
926         }
927     }

929     return (-1);
930 }

932 static void
933 procs_store(pid_t child)
934 {
935     int i;

937     i = procs_find((pid_t)(0));
938     if (i < 0) {
939         PERR(NOCHILDSLOT);
940         exit(1);
941     }
942     procs[i] = child;
943     n_procs++;
944 }

946 static int
947 procs_delete(pid_t child)
948 {
949     int i;

951     i = procs_find(child);
952     if (i < 0) {
953         return (0);
954     }
955     procs[i] = (pid_t)(0);
956     n_procs--;
957     return (1);
958 }

```

```

960 static pid_t
961 procs_waitpid(int blocking, int *stat_loc)
962 {
963     pid_t child;
964     int options;

966     if (n_procs == 0) {
967         errno = ECHILD;
968         return (-1);
969     }

971     options = (blocking) ? 0 : WNOHANG;

973     while ((child = waitpid(-1, stat_loc, options)) > 0) {
974         if (procs_delete(child)) {
975             break;
976         }
977     }

979     return (child);
980 }

982 static void
983 procs_wait(int blocking)
984 {
985     pid_t child;
986     int stat_loc;

988     while ((child = procs_waitpid(blocking || (n_procs >= MAXPROCS) ? 1 : 0,
989         if (WIFSIGNALED(stat_loc)) {
990             EMSG2(CHILDSIG, WTERMSIG(stat_loc));
991             exit(125);
992             /* NOTREACHED */
993         } else if ((WEXITSTATUS(stat_loc) & 0377) == 0377) {
994             EMSG(CHILDFAIL);
995             exit(124);
996             /* NOTREACHED */
997         } else {
998             exitstat |= WEXITSTATUS(stat_loc);
999         }
1000     }

1002     if (child == (pid_t)(-1) && errno != ECHILD) {
1003         EMSG(WAITFAIL);
1004         exit(122);
1005         /* NOTREACHED */
1006     }
1007 }
1008 #endif /* ! codereview */

1010 static void
1011 usage()
1012 {
1013     ermsg(_(USAGEMSG));
1014     OK = FALSE;
1015 }

1019 /*
1020 * parseargs():      modify the args
1021 *                   since the -e, -i and -l flags all take optional subarguments,
1022 *                   and getopt(3C) is clueless about this nonsense, we change the
1023 *                   our local argument count and strings to separate this out,
1024 *                   and make it easier to handle via getopt(3c).

```



```

1025 *
1026 *      -e      -> "-e ""
1027 *      -e3     -> "-e "3"
1028 *      -Estr   -> "-E "str"
1029 *      -i      -> "-i "{}"
1030 *      -irep   -> "-i "rep"
1031 *      -l      -> "-i "1"
1032 *      -l10    -> "-i "10"
1033 *
1034 *      since the -e, -i and -l flags all take optional subarguments,
1035 */
1036 static void
1037 parseargs(int ac, char **av)
1038 {
1039     int i;                /* current argument */
1040     int cflag;            /* 0 = not processing cmd arg */
1041
1042     if ((mav = malloc((ac * 2 + 1) * sizeof (char *))) == NULL) {
1043         PERR(MALLOCFAIL);
1044         exit(1);
1045     }
1046
1047     /* for each argument, see if we need to change things: */
1048     for (i = mac = cflag = 0; (av[i] != NULL) && i < ac; i++, mac++) {
1049         if ((mav[mac] = strdup(av[i])) == NULL) {
1050             PERR(MALLOCFAIL);
1051             exit(1);
1052         }
1053
1054         /* -- has been found or argument list is fully processed */
1055         if (cflag)
1056             continue;
1057
1058         /*
1059          * if we're doing special processing, and we've got a flag
1060          */
1061         else if ((av[i][0] == '-' && (av[i][1] != NULL)) {
1062             char *def;
1063
1064             switch (av[i][1]) {
1065             case 'e':
1066                 def = ""; /* -e with no arg turns off eof */
1067                 goto process_special;
1068             case 'i':
1069                 def = INSPAT_STR;
1070                 goto process_special;
1071             case 'l':
1072                 def = "1";
1073             process_special:
1074                 /*
1075                  * if there's no sub-option, we *must* add
1076                  * a default one. this is because xargs must
1077                  * be able to distinguish between a valid
1078                  * suboption, and a command name.
1079                  */
1080                 if (av[i][2] == NULL) {
1081                     mav[++mac] = strdup(def);
1082                 } else {
1083                     /* clear out our version: */
1084                     mav[mac][2] = NULL;
1085                     mav[++mac] = strdup(&av[i][2]);
1086                 }
1087                 if (mav[mac] == NULL) {
1088                     PERR(MALLOCFAIL);
1089                     exit(1);
1090                 }

```

```

1091         break;
1092
1093         /* flags with required subarguments: */
1094
1095         /*
1096          * there are two separate cases here. either the
1097          * flag can have the normal XCU4 handling
1098          * (of the form: -X subargument); or it can have
1099          * the old solaris 2.[0-4] handling (of the
1100          * form: -Xsubargument). in order to maintain
1101          * backwards compatibility, we must support the
1102          * latter case. we handle the latter possibility
1103          * first so both the old solaris way of handling
1104          * and the new XCU4 way of handling things are allowed.
1105          */
1106         case 'n': /* FALLTHROUGH */
1107         case 'p': /* FALLTHROUGH */
1108             #endif /* ! codereview */
1109         case 's': /* FALLTHROUGH */
1110         case 'E': /* FALLTHROUGH */
1111         case 'I': /* FALLTHROUGH */
1112         case 'L': /*
1113          * if the second character isn't null, then
1114          * the user has specified the old syntax.
1115          * we move the subargument into our
1116          * mod'd argument list.
1117          */
1118             if (av[i][2] != NULL) {
1119                 /* first clean things up: */
1120                 mav[mac][2] = NULL;
1121
1122                 /* now add the separation: */
1123                 ++mac; /* inc to next mod'd arg */
1124                 if ((mav[mac] = strdup(&av[i][2])) ==
1125                     NULL) {
1126                     PERR(MALLOCFAIL);
1127                     exit(1);
1128                 }
1129                 break;
1130             }
1131             i++;
1132             mac++;
1133
1134             if (av[i] == NULL) {
1135                 mav[mac] = NULL;
1136                 return;
1137             }
1138             if ((mav[mac] = strdup(av[i])) == NULL) {
1139                 PERR(MALLOCFAIL);
1140                 exit(1);
1141             }
1142             break;
1143
1144         /* flags */
1145         case 'p':
1146         case 't':
1147         case 'x':
1148         case '0':
1149             break;
1150
1151         case '-':
1152         default:
1153             /*
1154              * here we've hit the cmd argument. so
1155              * we'll stop special processing, as the

```

```
1157             * cmd may have a "-i" etc., argument,
1158             * and we don't want to add a "" to it.
1159             */
1160             cflag = 1;
1161             break;
1162         }
1163     } else if (i > 0) { /* if we're not the 1st arg */
1164         /*
1165          * if it's not a flag, then it *must* be the cmd.
1166          * set cflag, so we don't mishandle the -[eil] flags.
1167          */
1168         cflag = 1;
1169     }
1170 }

1172 mav[mac] = NULL;
1173 }
```



```

126 .na
127 \fB\fB-l[\fR\fInumber\fR\fB]\fR\fR
128 .ad
129 .RS 15n
130 (The letter ell.) This option is equivalent to \fB-L\fR \fInumber\fR. If
131 \fInumber\fR is omitted, \fB-l\fR is assumed. Option \fB-x\fR is forced on.
132 .RE

134 .sp
135 .ne 2
136 .na
137 \fB\fB-n\fR \fInumber\fR\fR
138 .ad
139 .RS 15n
140 Invokes \fIutility\fR using as many standard input arguments as possible, up to
141 \fInumber\fR (a positive decimal integer) arguments maximum. Fewer arguments
142 are used if:
143 .RS +4
144 .TP
145 .ie t \(\bu
146 .el o
147 The command line length accumulated exceeds the size specified by the \fB-s\fR
148 option (or \fB{LINE_MAX}\fR if there is no \fB-s\fR option), or
149 .RE
150 .RS +4
151 .TP
152 .ie t \(\bu
153 .el o
154 The last iteration has fewer than \fInumber\fR, but not zero, operands
155 remaining.
156 .RE
157 .RE

159 .sp
160 .ne 2
161 .na
162 \fB\fB-p\fR
163 .ad
164 .RS 15n
165 Prompt mode. The user is asked whether to execute \fIutility\fR at each
166 invocation. Trace mode (\fB-t\fR) is turned on to write the command instance to
167 be executed, followed by a prompt to standard error. An affirmative response
168 (specific to the user's locale) read from \fB/dev/tty\fR executes the command;
169 otherwise, that particular invocation of \fIutility\fR is skipped.
170 .RE

172 .sp
173 .ne 2
174 .na
175 \fB\fB-P\fR \fImaxprocs\fR\fR
176 .ad
177 .RS 15n
178 Invokes \fIutility\fR using at most \fImaxprocs\fR (a positive decimal integer)
179 parallel child processes.
180 #endif /* ! codereview */
181 .RE

183 .sp
184 .ne 2
185 .na
186 \fB\fB-s\fR \fIsize\fR\fR
187 .ad
188 .RS 15n
189 Invokes \fIutility\fR using as many standard input arguments as possible
190 yielding a command line length less than \fIsize\fR (a positive decimal
191 integer) bytes. Fewer arguments are used if:

```

```

192 .RS +4
193 .TP
194 .ie t \(\bu
195 .el o
196 The total number of arguments exceeds that specified by the \fB-n\fR option, or
197 .RE
198 .RS +4
199 .TP
200 .ie t \(\bu
201 .el o
202 The total number of lines exceeds that specified by the \fB-L\fR option, or
203 .RE
204 .RS +4
205 .TP
206 .ie t \(\bu
207 .el o
208 End of file is encountered on standard input before \fIsize\fR bytes are
209 accumulated.
210 .RE
211 Values of \fIsize\fR up to at least \fB{LINE_MAX}\fR bytes are supported,
212 provided that the constraints specified in DESCRIPTION are met. It is not
213 considered an error if a value larger than that supported by the implementation
214 or exceeding the constraints specified in DESCRIPTION is specified. \fBxargs\fR
215 uses the largest value it supports within the constraints.
216 .RE

218 .sp
219 .ne 2
220 .na
221 \fB\fB-t\fR\fR
222 .ad
223 .RS 6n
224 Enables trace mode. Each generated command line is written to standard error
225 just prior to invocation.
226 .RE

228 .sp
229 .ne 2
230 .na
231 \fB\fB-x\fR\fR
232 .ad
233 .RS 6n
234 Terminates if a command line containing \fInumber\fR arguments (see the
235 \fB-n\fR option above) or \fInumber\fR lines (see the \fB-L\fR option above)
236 does not fit in the implied or specified size (see the \fB-s\fR option above).
237 .RE

239 .sp
240 .ne 2
241 .na
242 \fB\fB-0\fR
243 .ad
244 .RS 6n
245 Null separator mode. Instead of using white space or new lines to
246 delimit arguments, zero bytes are used. This is suitable for use with
247 the -print0 argument to \fBfind\fR(1).
248 .RE

250 .SH OPERANDS
251 .sp
252 .LP
253 The following operands are supported:
254 .sp
255 .ne 2
256 .na
257 \fB\fIutility\fR\fR

```

```

258 .ad
259 .RS 12n
260 The name of the utility to be invoked, found by search path using the
261 \fBPATH\fR environment variable. (see \fBenviron\fR(5).) If \fIutility\fR is
262 omitted, the default is the \fBecho\fR(1) utility. If the \fIutility\fR operand
263 names any of the special built-in utilities in \fBshell_builtins\fR(1), the
264 results are undefined.
265 .RE

267 .sp
268 .ne 2
269 .na
270 \fB\fIargument\fR\fR
271 .ad
272 .RS 12n
273 An initial option or operand for the invocation of \fIutility\fR.
274 .RE

276 .SH USAGE
277 .sp
278 .LP
279 The \fB255\fR exit status allows a utility being used by \fBxargs\fR to tell
280 \fBxargs\fR to terminate if it knows no further invocations using the current
281 data stream succeeds. Thus, \fIutility\fR should explicitly \fBexit\fR with an
282 appropriate value to avoid accidentally returning with \fB255\fR.
283 .sp
284 .LP
285 Notice that input is parsed as lines. Blank characters separate arguments. If
286 \fBxargs\fR is used to bundle output of commands like \fBfind\fR \fBldir\fR
287 \fB-print\fR or \fBls\fR into commands to be executed, unexpected results are
288 likely if any filenames contain any blank characters or newline characters.
289 This can be fixed by using \fBfind\fR to call a script that converts each file
290 found into a quoted string that is then piped to \fBxargs\fR. Notice that the
291 quoting rules used by \fBxargs\fR are not the same as in the shell. They were
292 not made consistent here because existing applications depend on the current
293 rules and the shell syntax is not fully compatible with it. An easy rule that
294 can be used to transform any string into a quoted form that \fBxargs\fR
295 interprets correctly is to precede each character in the string with a
296 backslash (\fB\e\fR).
297 .sp
298 .LP
299 On implementations with a large value for \fB{ARG_MAX}\fR, \fBxargs\fR can
300 produce command lines longer than \fB{LINE_MAX}\fR. For invocation of
301 utilities, this is not a problem. If \fBxargs\fR is being used to create a text
302 file, users should explicitly set the maximum command line length with the
303 \fB-s\fR option.
304 .sp
305 .LP
306 The \fBxargs\fR utility returns exit status \fB127\fR if an error occurs so
307 that applications can distinguish "failure to find a utility" from "invoked
308 utility exited with an error indication." The value \fB127\fR was chosen
309 because it is not commonly used for other meanings; most utilities use small
310 values for "normal error conditions" and the values above \fB128\fR can be
311 confused with termination due to receipt of a signal. The value \fB126\fR was
312 chosen in a similar manner to indicate that the utility could be found, but not
313 invoked.
314 .SH EXAMPLES
315 .LP
316 \fBExample 1\fR Using the xargs command
317 .sp
318 .LP
319 The following example moves all files from directory \fB$1\fR to directory
320 \fB$2\fR, and echo each move command just before doing it:

322 .sp
323 .in +2

```

```

324 .nf
325 example% \fBls $1 | xargs -I {} -t mv $1/{ } $2/{ }\fR
326 .fi
327 .in -2
328 .sp

330 .sp
331 .LP
332 The following command combines the output of the parenthesised commands onto
333 one line, which is then written to the end of file \fBlog\fR:

335 .sp
336 .in +2
337 .nf
338 example% \fB(logname; date; printf "%s\n" "$0 $*") | xargs >>log\fR
339 .fi
340 .in -2
341 .sp

343 .sp
344 .LP
345 The following command invokes \fBdiff\fR with successive pairs of arguments
346 originally typed as command line arguments (assuming there are no embedded
347 blank characters in the elements of the original argument list):

349 .sp
350 .in +2
351 .nf
352 example% \fBprintf "%s\n" "$*" | xargs -n 2 -x diff\fR
353 .fi
354 .in -2
355 .sp

357 .sp
358 .LP
359 The user is asked which files in the current directory are to be archived. The
360 files are archived into \fBarch\fR \fB;\fR a, one at a time, or b, many at a
361 time:

363 .sp
364 .in +2
365 .nf
366 example% \fBls | xargs -p -L 1 ar -r arch
367 ls | xargs -p -L 1 | xargs ar -r arch\fR
368 .fi
369 .in -2
370 .sp

372 .sp
373 .LP
374 The following executes with successive pairs of arguments originally typed as
375 command line arguments:

377 .sp
378 .in +2
379 .nf
380 example% \fBecho $* | xargs -n 2 diff\fR
381 .fi
382 .in -2
383 .sp

385 .SH ENVIRONMENT VARIABLES
386 .sp
387 .LP
388 See \fBenviron\fR(5) for descriptions of the following environment variables
389 that affect the execution of \fBxargs\fR: \fBBLANG\fR, \fBBLC_ALL\fR,

```

```

390 \fBLC_COLLATE\fR, \fBLC_CTYPE\fR, \fBLC_MESSAGES\fR, and \fBNLSPATH\fR.
391 .sp
392 .ne 2
393 .na
394 \fB\fBPATH\fR\fR
395 .ad
396 .RS 8n
397 Determine the location of \fIutility\fR.
398 .RE

400 .sp
401 .LP
402 Affirmative responses are processed using the extended regular expression
403 defined for the \fByesexpr\fR keyword in the \fBLC_MESSAGES\fR category of the
404 user's locale. The locale specified in the \fBLC_COLLATE\fR category defines
405 the behavior of ranges, equivalence classes, and multi-character collating
406 elements used in the expression defined for \fByesexpr\fR. The locale specified
407 in \fBLC_CTYPE\fR determines the locale for interpretation of sequences of
408 bytes of text data a characters, the behavior of character classes used in the
409 expression defined for the \fByesexpr\fR. See \fBlocale\fR(5).
410 .SH EXIT STATUS
411 .sp
412 .LP
413 The following exit values are returned:
414 .sp
415 .ne 2
416 .na
417 \fB\fB0\fR\fR
418 .ad
419 .RS 12n
420 All invocations of \fIutility\fR returned exit status \fB0\fR.
421 .RE

423 .sp
424 .ne 2
425 .na
426 \fB\fB1\mi125\fR\fR
427 .ad
428 .RS 12n
429 A command line meeting the specified requirements could not be assembled, one
430 or more of the invocations of \fIutility\fR returned a non-zero exit status, or
431 some other error occurred.
432 .RE

434 .sp
435 .ne 2
436 .na
437 \fB\fB126\fR\fR
438 .ad
439 .RS 12n
440 The utility specified by \fIutility\fR was found but could not be invoked.
441 .RE

443 .sp
444 .ne 2
445 .na
446 \fB\fB127\fR\fR
447 .ad
448 .RS 12n
449 The utility specified by \fIutility\fR could not be found.
450 .RE

452 .sp
453 .LP
454 If a command line meeting the specified requirements cannot be assembled, the
455 utility cannot be invoked, an invocation of the utility is terminated by a

```

```

456 signal, or an invocation of the utility exits with exit status \fB255\fR, the
457 \fBxargs\fR utility writes a diagnostic message and exit without processing any
458 remaining input.
459 .SH ATTRIBUTES
460 .sp
461 .LP
462 See \fBattributes\fR(5) for descriptions of the following attributes:
463 .sp

465 .sp
466 .TS
467 box;
468 c | c
469 l | l .
470 ATTRIBUTE TYPE      ATTRIBUTE VALUE
471 _
472 CSI      Enabled
473 _
474 Interface Stability      Standard
475 .TE

477 .SH SEE ALSO
478 .sp
479 .LP
480 \fBBecho\fR(1), \fBshell_builtins\fR(1), \fBattributes\fR(5), \fBenviron\fR(5),
481 \fBstandards\fR(5)

```