                        CCA ICCC WEATHER DEMO

   THE WEATHER DEMO ORIGINALLY DEVELOPED FOR THE ICCC SHOW IS NOW
   GENERALLY AVAILABLE. IT IS A VERY NICE NETWORK DEMO SINCE IT REALLY
   CAN/DOES RUN ON TWO MACHINES.

   THIS DEMO IS TWO INDEPENDENT PROGRAMS. TOGETHER THEY TRY TO ANSWER
   ENGLISH QUESTIONS ABOUT THE WEATHER. FOR EXAMPLE, 'DID IT RAIN IN
   BOSTON ON OCTOBER 7?'. THE NATURAL LANGUAGE PROGRAM IS WRITTEN IN
   LISP. IT IS A MODIFICATION OF A SYSTEM DEVELOPED BY T. WINOGRAD AT
   MIT. THE OTHER HALF IS A PRIMITIVE DATACOMPUTER. THE DATACOMPUTER
   PROVIDES NETWORK DATAMANAGEMENT FACILITIES AND WILL INCORPORATE THE
   TRILLION-BIT UNICON CURRENTLY BEING INSTALLED AT AMES.

   THE TWO PROGRAMS COMMUNICATE VIA DATALANGUAGE USING THE NET.

   'WEATHER' NOW USES HOST '37' AS THE LOCATION OF THE DATACOMPUTER.
   THIS FIXES THE OLD 'CCA' NOT IN THE HOST TABLES BUG.

   KEEP IN MIND THAT IS AN EVOLVING SYSTEM, WE PLAN TO IMPROVE 'WEATHER'
   AND THE DATACOMPUTER THAT IT USES, SO THINGS THAT WORK ONE DAY MAY
   GET BROKEN THE NEXT.

AVAILABILITY

   'WEATHER' (THE LISP HALF) IS A REAL MACHINE USER - BOTH CPU AND CORE.
   WE WOULD OBVIOUSLY PREFER THAT YOU RUN IT ON YOUR MACHINE. IF WE ARE
   REALLY EMPTY (WHICH FREQUENTLY HAPPENS AT 3 AM) FEEL FREE TO RUN IT
   ON OUR MACHINE. SINCE OUR SYSTEM HAS NO DRUM, AND WEATHER TAKES ALL
   OF CORE, PLEASE CHECK WITH SOMEBODY AT CCA BEFORE RUNNING WEATHER IF
   WE ARE DOING ANYTHING AT ALL. IT IS REALLY A MAJOR LOAD ON OUR
   MACHINE.

   SINCE THE DATACOMPUTER IS LESS OF A LOAD, THE SYSTEM WILL BE MORE
   AVAILABLE IF YOU CAN RUN 'WEATHER' ON YOUR MACHINE. WE WILL TRY TO
   RUN THE DATACOMPUTER DURING NON-PRIME TIME. IT CAN BE THOUGHT OF AS
   THE SERVER HALF OF THIS COMBINATION. PRIME TIME HERE IS 1 TO 5 PM
   EASTERN TIME, BUT THE MORNING IS FREQUENTLY QUITE BUSY TOO. PLEASE
   DO NOT USE THE DATACOMPUTER DURING THIS TIME. IT WILL SAVE US THE
   EFFORT OF TURNING IT OFF. CURRENTLY, ONLY ONE USER AT A TIME CAN RUN
   - A LIMITATION OF THE DATACOMPUTER THAT WILL GO AWAY SOMEDAY.

IF YOU WANT TO RUN A DEMO (NOT JUST PLAY AROUND) WE WILL PROBABLY BE
ABLE TO COOPERATE. PLEASE CALL TO MAKE ARRANGEMENTS WELL IN ADVANCE.
WE STRONGLY RECOMMEND THAT YOU DO NOT TRY TO GIVE AN IMPORTANT DEMO
UNLESS YOU HAVE USED BOTH VERSIONS OF THE SYSTEMS THAT WILL BE USED
FOR THE DEMO. EVEN WE GET IN TROUBLE WHEN WE FORGET TO REHEARSE
THINGS A BIT.

WEDNESDAY MORNING IS SCHEDULED PREVENTIVE MAINTENANCE TIME. OUR
MACHINE WILL NOT BE UP BETWWEN 9 AND 12, BUT IT FREQUENTLY IS DOWN
PAST NOON.

MORE INFORMATION

FOR THE LATEST INFORMATION, LOOK AT *.TXT IN <WDB-DEMO> AT CCA. IN
PARTICULAR, MESSAGE.TXT SHOULD BE A QUICK SUMMARY OF ANY OTHER
INTERESTING TXT FILES. THE LATEST VERSION OF THIS FILE (IN RUNOFF
FORMAT) IS <HGM-MEMOS>ANNOUNCE.RNO.

THE LATEST VERSION OF THE LISP PROGRAM WILL BE IN <WDB-DEMO> AS
WEATHER.SAV. NORMAL FTP CAN BE USED TO COPY IT - IT IS 400+ PAGES.
IT TAKES 6 MIN TO TRANSFER IT TO OURSELVES (1.5 MIN OF CPU, 22KBAUD)
WHEN THE MACHINE IS VERY LIGHLY LOADED. BE SURE TO SET THE BYTE SIZE
TO 36 ('TENEX' COMMAND TO FTP) SINCE THIS IS A SAVE FILE.

PROBLEMS

IF YOU FIND ANY PROBLEMS WITH LISP HALF, PLEASE USE SNDMSG TO SEND
THINGS TO JEFF HILL - HE IS JMH@CCA. WE ARE INTERESTED IN EXAMPLES
WHICH LOOK LIKE THEY SHOULD WORK BUT DON'T. IF YOU HAVE ANY COMMENTS
ABOUT GENERAL OPERATIONS GLITCHES - LIKE THE DATACOMPUTER IS NOT UP -
BITCH TO ME - HGM@CCA. CCA IS HOST 31 DECIMAL, 37 OCTAL.

USING THE SYSTEM

AFTER STARTING WEATHER, THE SYSTEM WILL INITIALIZE ITSELF, TYPE OUT
SOME MESSAGES, AND ASK FOR THE JULIAN DATE (REALLY THE DAY NUMBER OF
THE YEAR OF 'TODAY'). 300 IS THE BEST ANSWER. THAT MAKES 'TODAY' BE
OCT 26, THE LAST DAY OF THE ICCC DEMO, AND ONE DAY PAST THE LATEST
DATA. IF YOU TYPE SOME OTHER NUMBER, 'YESTERDAY', AND 'LAST WEEK'
WON'T WORK AS WELL. THE SYSTEM PROMPTS WITH 'READY' WHEN IT IS READY
TO ACCEPT A SENTENCE. BE SURE TO END THE SENTENCE WITH THE PROPER
PUNCTUATION. SENTENCES MAY BE EDITED WITH:

    ^A (CONTROL A) -- DELETES LAST CHARACTER OR WORD TYPED

    ^R (CONTROL R) -- RETYPES SENTENCE SO FAR

    ^Q (CONTROL Q) -- DELETES ENTIRE SENTENCE

     ^X (CONTROL X) -- SAME AS CONTROL Q

   IN ADDITION, TWO CONTROL CHARACTERS CAN BE USED DURING THE PROCESSING
   OF SENTENCES, ^E (CONTROL E) ABORTS THE PROCESSING OF THE SENTENCE
   AND RETURNS TO THE READY STATE. ^D (CONTROL D) ABORTS THE WHOLE
   SYSTEM, AND CAUSES IT TO RE-INITIALIZE ITSELF.

   DURING SENTENCE INPUT, THE SYSTEM ATTEMPTS TO LOOK UP EACH WORD AS IT
   IS TYPED, AND DO SPELLING CORRECTION. HOWEVER, SINCE THE NAMES OF
   ALL THE CITIES ARE KEPT AT THE DATACOMPUTER, THE SYSTEM ACCEPTS ALL
   UNKNOWN WORDS AS POSSIBLE PLACE NAMES TO BE CHECKED LATER ON IN THE
   PROCESSING. IF THE SPELLING CORRECTOR CORRECTS ONE OF YOUR CITY
   NAMES TO A VOCABULARY WORD, YOU CAN ERASE THE WORD AND RETYPE IT
   PRECEEDED BY AN EQUAL SIGN (=). THIS FORCES THE SYSTEM TO ACCEPT THE
   WORD AS A PROPER NOUN.

   THE SYSTEM RETURNS TO THE READY STATE AFTER ANSWERING THE QUESTION.
   TO TERMINATE A SESSION YOU MAY TYPE 'GOODBYE.', WHICH WILL LEAVE YOU
   BACK AT THE TENEX COMMAND LEVEL.

   IF THE SYSTEM CANNOT UNDERSTAND YOUR REQUEST, IT TRIES TO GIVE YOU
   SOME EXPLANATORY MESSAGE. HOWEVER, IT'S NOT VERY GOOD AT EXPLAINING
   ITS BEHAVIOR SO YOU WILL USUALLY GET THE DEFAULT MESSAGE 'SORRY, I
   COULDN'T UNDERSTAND THAT REQUEST.' IF YOU REALLY WANT TO KNOW WHY,
   SEND US A TRANSCRIPTIPT OF YOUR SESSION.

RESTRICTIONS - HINTS

   THE SYSTEM IS NOT LINGUISTICALLY SOPHISTICATED. IT PERFORMS PASSABLY
   ON SENTENCES WHICH WOULD BE CONSIDERED SIMPLE REQUESTS FOR
   INFORMATION. IT DOES NOT REALLY UNDERSTAND THE 'WEATHER', RATHER IT
   UNDERSTANDS HOW TO COMPOSE RETRIEVAL REQUESTS FOR WEATHER DATA IN ITS
   DATA BASE.

   A LARGE SET OF EXAMPLE SENTENCES CAN BE FOUND IN <WDB-
   DEMO>SENTENCES.TXT. THOSE SENTENCES PRECEDE BY A BACKSLASH (\) DO
   NOT WORK. AN EXPLANATORY COMMENT, PRECEEDED BY A SEMI-COLON (;),
   EXPLAINS THE REASONS FOR FAILURE. SOME OF THE OTHERS MAY NOT WORK
   EITHER.

   THE MAJOR THINGS THAT WE KNOW DON'T WORK ARE:

      1. CONJUNCTION

      2. NEGATION

      3. COMPLEX QUANTIFICATION

4. SOPHISTICATED PRONOUN REFERENCE

5. QUESTIONS NOT (!!) INVOLVING WEATHER,E.G.,

    'WHAT DAY WAS YESTERDAY?'

    'WHAT CITIES ARE IN MASSACHUSETTS?'

'IT' IS NOT WORKING TOO WELL THESE DAYS. IF YOU GET IN TROUBLE, TRY
SOMETHING WITHOU 'IT' THAT IS KNOWN TO WORK. 'WHAT WAS THE
TEMPERATURE IN BOSTON YESTERDAY?' IS BETTER THAN MOST - WE ARE QUITE
APT TO TEST THAT PHRASE.

THE SYSTEM DOES NOT KNOW ABOUT THE WORD 'SNOW'. THERE IS NO WAY TO
LIST ITS VOCABULARY. HOWEVER, THE SENTENCES IN SENTENCES.TXT SHOULD
GIVE YOU A GOOD IDEA OF THE WORDS THE SYSTEM UNDESTANDS.

THE AVERAGE TIME FOR UNDERSTANDING AND ANSWERING A QUESTION IS ABOUT
1 MINUTE. IF THE SYSTEM TAKES MUCH LONGER THAN THIS YOU SHOULD GET
SUSPICIOUS. IF YOU ASK A VERY BIG QUESTION, LIKE 'WHAT WAS THE
HOTTEST PLACE IN CANADA?', THE LISP SYSTEM MAY GO INTO A GC (GARBAGE
COLLENTION) LOOP. RESTARTING ALL OVER (WITH ^D) WILL BE OK, BUT THE
DATACOMPUTER MAY TAKE A LONG TIME TO RECOVER - IT DOESN'T (YET)
NOTICE THAT ANYTHING HAS GONE WRONG UNTIL IT ASKS FOR MORE INPUT.

DATA

THE DATA FROM OCT 1 THRU 16 IS ALMOST COMPLETE. BETWEEN OCT 17 TO 25
IT IS FREQUENTLY SKETCHY. WE PLAN TO RELOAD EVERYTHING SOMETIME
USING SEPTEMBER'S DATA, SO WE WILL HAVE A COMPLETE MONTH OF CLEAN
DATA. NOTE: THIS DATA HAS BEEN EXTENSIVELY SUMMARIZED - IT IS
PROBABLY NOT USEFUL FOR PEOPLE INTERESTED IN WEATHER.

THE OCTOBER DATA IS ABOUT 2400 PAGES OF REAL DATA, PLUS ABOUT 200
MORE FOR THE INVERSION. THERE ARE ABOUT 4000 CITIES. THE CITY NAMES
HAVE BEEN CHANGED TO THE LOCAL CITY WHENEVER POSSIBLE. I.E. 'BOSTON'
IS 'BOSTON' NOT 'LOGAN'.

'ALASKA' GOT LOST SOMEHOW. DON'T BOTHER ASKING FOR IT.

SAMPLE DATALANGUAGE

THESE EXAMPLES ON THIS AND THE FOLLOWING PAGES WORK ON A VERY OLD
VERSION OF THE DATACOMPUTER. THEY ARE NOT THE FINAL FORM OF
DATALANGUAGE THAT SHOULD BA AVAILABLE SOON.

THE FOLLOWING QUESTION TO 'WEATHER' GENERATED TWO REQUESTS TO THE
DATACOMPUTER. THE FIRST USES THE CHECK PORT TO VERIFY THAT 'BOSTON'
IS A VALID CITY NAME, AND THE SECOND USES THE ANSWER PORT TO RETRIEVE
THE DATA. THE ACTUAL REQUESTS THAT ARE SENT TO THE DATACOMPUTER DO
NOT HAVE THE CARRIAGE RETURNS THAT MAKE THEM READABLE - THEY ARE SENT
AS ONE VERY LONG LINE. THE '^Z' IS THE END OF DATA MARKER.

TO WEATHER

HAS IT RAINED IN BOSTON RECENTLY?

TO THE DATACOMPUTER:

```
OR CHECK.PLACE,STATION WITH CITY EQ 'BOSTON'
PLACE.LP1 = '('
PLACE.CITY = STATION,CITY
PLACE.RP1 = ')'
PLACE.BSN = STATION.BSN
PLACE.LP2 = '('
PLACE.REGION = STATION.REGION
PLACE.RP2 = ')'
END
```

FROM THE DATACOMPUTER:

```
(BOSTON                  )725090(MASSACHUSETTS           )
^Z
```

TO THE DATACOMPUTER:

```
FOR THE STATION WITH (REGION EQ 'MASSACHUSETTS' AND CITY EQ 'BOSTON')
 FOR ANSWER.ANS OBSERVATION WITH GE '273' AND DATE LE '300'
        ANS.LP1 = '('
        ANS.CITY = STATION.CITY
        ANS.RP1 =')'
        ANS.DATE = OBSERVATION.DATE
        ANS.LP2 = ')'
        ANS.DATA = PRECIP
        ANS.DATA1 = '     '
        ANS.DATA2 = '     '
        ANS.RP2 = ')'
  END
 END
```

    AND BACK FROM THE DATACOMPUTER:

    (BOSTON   )293(   )
    (BOSTON   )294(0   )
    (BOSTON   )295(   )
    (BOSTON   )296(10   )
    (BOSTON   )297(0   )
    (BOSTON   )298(0   )
    (BOSTON   )299(0   )
    (BOSTON   )300(   )
    ^Z

    AND THE FINAL ANSWER FROM WEATHER WAS 'YES.'

DESCRIPTIONS

    THE FOLLOWING IS THE DESCRIPTION OF THE FILE AS IT IS STORED IN THE
    DATACOMPUTER.

    CREATE WEATHER FILE LIST
      STATION STRUCT
            BSN   STR(6),   I=D
            CITY   STR(22),   I=D
            REGION   STR(22),   I=D
            WORLD   STR(22)
            OBS   LIST (31)
             OBSERVATION   STRUCT
              DATE   STR(3)
              TEMPERATURE   STRUCT
                         MIN STR(4)
                         MAX STR(4)
              END TEMPERATURE   STRUCT
              PRECIP   STR(4)
              WINDS   STRUCT
                         SPEED   STR(4)
                         GUSTS   STR(4)
                         DIRECTION   STR(4)
              END WINDS STRUCT
              VISIBILITY   STR(4)
              CLOUDS   STR(4)
              GENERAL STR(4)
              PRESSURE
              END OBSERVATION STRUCT
            END STATION STRUCT
      END WEATHER FILE

```
   THE FOLLOWING PORT DESCRIPTIONS ARE USED BY WEATHER TO RETRIVE DATA
   FROM THE DATACOMPUTER.

   CREATE CHECK PORT LIST
                   PLACE STRUCT
                   LP1 STR (1)
                   CITY STR (22)
                   RP1 STR (1)
                   BSN STR (6)
                   LP2 STR (1)
                   REGION STR (22)
                   RP2 STR (1)
                   END PLACE STRUCT
   END CREATE CHECK

   CREATE ANSWER PORT LIST
                   ANS STRUCT
                   LP1 STR (1)
                   CITY STR (22)
                   RP1 STR (1)
                   DATE STR (3)
                   LP2 STR (1)
                   DATA STR (4)
                   DATA1 STR (4)
                   DATA2 STR (4)
                   RP2 STR (1)
                   END ANS STRUCT
   END CREATE ANSWER
```

SAMPLE RECORD

   THE FOLLOWING DATA IS SAMPLE RECORD - THE FIRST ONE IN THE FILE. IN
   THE DATACOMPUTER IT IS STORED WITHOUT THE CARRIAGE RETURNS NEEDED TO
   MAKE IT PRINT NEATLY.

```
   010100ANDOYA                      NORWAY                    EUROPE
   275
   276 278 279    0     2     0    12    9    758883023
   277 279 283    4    23    52    11    8    788882982
   278
   279 274 281    0    23    42    12    9    788882976
   280 268 270    4    14     0     9   19    627772983
   281 276 283    0    28    69     7    9    688882986
   282 269 273    4    10     0     9   22    727772975
   283 278 284    0    13    37     9    7    766882972
   284            0    28     0     5    1    80006   0
   285 276 280    0     9     0     3   12    666682936
   286 273 275    0    15     0     3    7    877782960
   287 278 282    0    13    28    11   14    626662981
   288 269 271   28    30     0     6    7    888882908
   289 275 281    0    21    45     7    9    688882976
   290 268 269    0    20     0     8   22    622782960
   291 283 283    0    11     0     7   47    100013002
   292 274 274    0    11     0     1   47    500012996
   293
   294 272 272    0     5     0     7   47    400012949
   295
   296
   297 273 273    0     0     0     1    6    80007   0
   298 273 273    0     0     0     1    2    80007   0
   299 275 275    0     6     0     7   31    30000   0
   300
   301
   302
   303
   304
   305
```


        [This RFC was put into machine readable form for entry]
      [into the online RFC archives by Helene Morin, Via Genie, 12/99]