Sorting a Bajillion Variables: When SORTC and SORTN Subroutines Have Stopped Satisfying, User-Defined PROC FCMP Subroutines Can Leverage the Hash Object to Reorder Limitless Arrays

Troy Martin Hughes

ABSTRACT

The SORTC and SORTN built-in SAS® subroutines sort character and numeric data, respectively, and are sometimes referred to as "horizontal sorts" because they sort variables rather than observations. That is, all elements within a SORTC or SORTN sort must be maintained in a single observation. A limitation of SORTC and SORTN is their inability to sort more than 800 variables when called inside the FCMP procedure. To overcome this disagreeable, arbitrary threshold, user-defined subroutines can be engineered that leverage the hash object to sort limitless variables. The hash object orders values that are ingested into it using the ORDERED argument, which can specify either ASCENDING or DESCENDING. This text demonstrates three failure patterns that occur when the OF operator specifies an array inside the FCMP procedure, which affect both character and numeric arrays, and which cause all built-in functions and subroutines to fail with runtime errors.

CONTRASTING VERTICAL AND HORIZONTAL SORTS

Within Base SAS, the SORT procedure and the SQL procedure's ORDER BY clause are the primary methods to perform a "vertical" sort—that is, to sort observations within a data set by one or more variables. The hash object can also sort observations, and can do so both inside the DATA step as well as inside the FCMP procedure.

The U.S. Constitution, the text of which is included in Appendix A and saved to Constitution.txt, provides an excellent text to demonstrate SAS sorting. The following DATA step ingests the text into Const:

```
%let loc=c:\sas\;                                    * USER MUST CHANGE LOCATION *;
filename f "&loc.constitution.txt";
data const (keep=word);
   infile f end=done missover;
   length raw word $20;
   raw='X';
   do while(^missing(raw));
      input raw $ @;
      word=lowcase(compress(raw,,'ka'));
      if ^missing(word) then output;
      end;
run;
```

The COMPRESS function removes all punctuation, and each of the 4,422 lowercase words are saved to a separate observation. The SORT procedure performs a vertical sort, and reorders the observations by the Word variable:

```
proc sort data=const out=const_sorted;
   by word;
run;
```

The DATA step prints the first and last observations:

```
data _null_;
   set const_sorted end=done;
   by word;
   if _n_=1 or done then put _n_ word;
run;
```

The first word is "a" and the last word is "yeas"—as in "and the Yeas and Nays of the Members of either House":

```
1 a
4422 yeas
NOTE: There were 4422 observations read from the data set WORK.CONST_SORTED.
```

A "horizontal" sort, by contrast, sorts the values within a single observation. For example, given a range of variables from Word1 to Word4422, the Word1 variable would be initialized to "a" and the Word4422 variable would be initialized to "yeas." Within Base SAS, the SORTC subroutine sorts character variables, and the SORTN subroutine sorts numeric variables.

To demonstrate a horizontal sort, the TRANSPOSE procedure first swaps observations and variables—that is, each observation is initialized into an incrementally higher variable, from Word1 to Word4422:

```
proc transpose data=const out=const_wide prefix=word;
   var word;
run;
```

Prior to the sort, the variables are ordered as they appear in the text of the U.S. Constitution:

```
data _null_;
   set const_wide;
   put word1-word7;
run;
```

Thus, the DATA step prints the following first seven words:

```
we the people of the united states
NOTE: There were 1 observations read from the data set WORK.CONST_WIDE.
```

The SORTC subroutine performs a horizontal sort to reorder the variables:

```
data const_wide_sorted;
   set const_wide (drop=_name_);
   array words[*] $ word:;
   call sortc(of words[*]);
   put word1-word7;
run;
```

Thereafter, the log demonstrates that A's now occupy the first several positions in the Words array—because the letter "a" appears 52 times in the U.S. Constitution:

```
a a a a a a a
NOTE: There were 1 observations read from the data set WORK.CONST_WIDE_SORTED.
```

With horizontal sorting differentiated and demonstrated, the next sections introduce limitations in SORTC and SORTN inside the FCMP procedure, as well as a hash-based workaround to certain issues.

THREE INDEPENDENT FAILURES OF SORTC AND SORTN SUBROUTINES INSIDE FCMP

The first failure occurs when attempting to leverage the OF operator on an array passed (by value or by reference) to a user-defined function or subroutine. That is, any numeric or character array that is declared by a FUNCTION or SUBROUTINE statement cannot be specified by a subsequent built-in function or subroutine that uses OF.

For example, extracting the SORTC subroutine from the previous DATA step, and placing it inside a functionally equivalent user-defined subroutine will fail:

```
proc fcmp outlib=work.funcs.sort;
   subroutine reorder(var[*] $);
      outargs var;
      call sortc(of var[*]);
      endsub;
quit;
```

The subroutine fails to compile:

```
1462  proc fcmp outlib=work.funcs.sort;
1463     subroutine reorder(var[*] $);
```

```
1464        * outargs var;
1465        call sortc(of var[*]);
ERROR: The OF operator isn't allowed on an ARRAY with a dynamic size.
1466        endsub;
1467  quit;

NOTE: The SAS System stopped processing this step because of errors.
```

This failure occurs whether declaring a character or numeric array, whether declaring the array call method as *by reference* (i.e., using the OUTARGS statement) or *by value* (i.e., without OUTARGS), and whether attempting to call the SORTC or SORTN subroutine.

A second attempt to sort an array inside the FCMP procedure might instead decompose a long character variable into its constituent elements (i.e., array variables) inside FCMP—that is *during* the function call, not *prior to* the function call.

For example, consider the need to utilize SORTC inside a user-defined function on an array declared (using the ARRAY statement) inside that function. First, this functionality can be demonstrated successfully inside the DATA step. The following DATA step builds the Text data set, in which Long_text contains the 4,422 words of the U.S. Constitution:

```
data text (drop=word);
   set const end=done;
   length long_text $30000;
   retain long_text '';
   long_text=catx(' ',long_text,word);
   if done;
run;
```

To perform a horizontal sort of the Long_text variable, the Words array is declared and incrementally initialized, after which it is sorted using the SORTC subroutine:

```
data _null_ (drop=long_text cnt elements);
   set text;
   length elements 8 new_long_text $30000;
   array words[5000] $20;
   * initialize and reorder array;
   elements=countw(long_text);
   do cnt=1 to elements;
      words[cnt]=scan(long_text,cnt);
      end;
   call sortc(of words[*]);
   * shift array elements left;
   do cnt=1 to 5000;
      if cnt<=elements then words[cnt]=words[5000-elements+cnt];
      else words[cnt]='';
      end;
   * print first 7 words alphabetically;
   put words1-words7;
   new_long_text=catx(' ',of words[*]);
run;
```

When SORTC performs the horizontal sort, missing values are placed before all other words, and because the array contains more elements (5,000) than the actual number of words (4,422), SORTC initializes the first 578 variables (in the array) to missing values. Thus, a second loop is required to shift non-missing values to the left within the array so that Words1 is filled with "a" rather than a missing value.

Finally, the New_long_text variable (not shown) is initialized to the sorted list of words. The first seven elements of the array are printed to the log using the PUT statement:

```
a a a a a a a
```

This somewhat convoluted logic can be extracted from the DATA step and emplaced within a user-defined function. The creation of a user-defined function improves reusability (of the callable functionality), and

improves readability of the DATA step because the convoluted logic has been encapsulated inside the function.

For example, a function can be defined that is passed a character variable containing numerous words, and which returns the sorted representation of those words. The REORDER function includes the logic from the previous DATA step:

```
proc fcmp outlib=work.funcs.sort;
   function reorder(string $) $;
      length elements 8 ordered $30000;
      array words[5000] $20;
      * initialize and reorder array;
      elements=countw(string);
      do cnt=1 to elements;
         words[cnt]=scan(string,cnt);
         end;
      call sortc(of words[*]);
      * shift array elements left;
      do cnt=1 to 5000;
         if cnt<=elements then words[cnt]=words[5000-elements+cnt];
         else words[cnt]='';
         end;
      ordered=catx(' ',of words[*]);
      return(ordered);
      endfunc;
quit;
```

The function should work, but the following DATA step demonstrates the second failure of SORTC inside the FCMP procedure:

```
options cmplib=work.funcs;

data _null_;
   set text;
   length alphabetized $30000;
   alphabetized=reorder(long_text);
run;
```

The log shows the unfortunate limitations of SORTC (and SORTN) when called inside the FCMP procedure:

```
1308  data _null_;
1309     set text;
1310     length alphabetized $30000;
1311     alphabetized=reorder(long_text);
1312  run;
```

ERROR: Function or Subroutine SORTC_CALL expects 5000 arguments, this exceeds the
maximum number of arguments allowed, 800.
NOTE: The SAS System stopped processing this step because of errors.

Thus, SORTC is failing because of the arbitrary limitation that only 800 variables can be sorted.

One final attempt to sort an array inside the FCMP procedure leverages the NOSYMBOLS option of the ARRAY statement. Declaring an array using NOSYMBOLS inside the FCMP procedure is akin to declaring a temporary array using the _TEMPORARY_ option inside a DATA step.

The dynamic array is subsequently resized to the total number of elements (i.e., words) using the DYNAMIC_ARRAY built-in subroutine:

```
proc fcmp outlib=work.funcs.sort;
   function reorder(string $) $;
      length elements 8 ordered $30000;
      array words[1] $20 / nosymbols;
      elements=countw(string);
      call dynamic_array(words, elements);
      * initialize and reorder array;
      do cnt=1 to elements;
```

```
            words[cnt]=scan(string,cnt);
            end;
        call sortc(of words[*]);
        ordered=catx(' ',of words[*]);
        return(ordered);
        endfunc;
quit;
```

Dynamic arrays can be powerful at times because they reduce inefficiency caused by unused elements, such as the arbitrary declaration of a 5000-element array to hold only 4,422 words. However, dynamic arrays are also extremely limited because they are incompatible with the OF statement, and thus cannot be leveraged by built-in functions:

```
1410  proc fcmp outlib=work.funcs.sort;
1411     function reorder(string $) $;
1412        length elements 8 ordered $30000;
1413        array words[1] $20 / nosymbols;
1414        elements=countw(string);
1415        call dynamic_array(words, elements);
1416        * initialize and reorder array;
1417        do cnt=1 to elements;
1418            words[cnt]=scan(string,cnt);
1419            end;
1420        call sortc(of words[*]);
ERROR: Argument number 1 to the CALL subroutine SORTC must be a variable.
1421        ordered=catx(' ',of words[*]);
1422        return(ordered);
1423        endfunc;
1424  quit;
```

NOTE: The SAS System stopped processing this step because of errors.

Unfortunately, these three limitations of arrays within the FCMP procedure join forces to thwart the ability to sort using either SORTC or SORTN. Fortunately, the hash object can save the day, as demonstrated in the next two sections.

PERFORMING A HORIZONTAL HASH SORT OF A CHARACTER ARRAY

To recap, a DATA step solution to horizontal sorting can be achieved by calling the SORTC built-in subroutine:

```
data const_wide_sorted;
   set const_wide (drop=_name_);
   array words[*] $ word:;
   call sortc(of words[*]);
run;
```

However, an equivalent sort inside the FCMP procedure requires uploading the array into an ordered hash object, and subsequently overwriting the array by iterating the hash object:

```
proc fcmp outlib=work.funcs.sort;
   subroutine reorder(arr[*] $);
      outargs arr;
      length key $20 num cnt 8;
      declare hash h(ordered: 'ascending');
      declare hiter iter('h');
      rc=h.defineKey('key');
      rc=h.defineData('key','cnt');
      rc=h.defineDone();
      do num=1 to dim(arr);
         key=arr[num];
         * increment a key that has been seen;
         if h.find()=0 then do;
            cnt+1;
            rc=h.replace();
            end;
```

```
          * initialize first occurrence of key;
          else do;
             cnt=1;
             rc=h.add();
             end;
          end;
       * build ordered array;
       num=1;
       do while(iter.next()=0);
          do tot=1 to cnt;
             arr[num]=key;
             num+1;
             end;
          end;
       endsub;
quit;
```

The REORDER user-defined subroutine is called in the following DATA step, and (finally) successfully reorders the longer-than-800-element array:

```
data const_wide_sorted;
   set const_wide (drop=_name_);
   array words[*] $ word:;
   call reorder(of words);
run;
```

Also note that the flexibility and scalability of the hash object foregoes having to instantiate an arbitrarily sized array, such as the 5000-element array previously declared.

The previous REORDER definition separates the addition of a key (to the hash object) and the subsequent recurrence of a key, and thus separately relies on the ADD hash method and the REPLACE hash method, respectively. However, because the default behavior of REPLACE adds a key that is not found in the hash object, this logic can be simplified.

A functionally equivalent definition of REORDER now relies solely on REPLACE, and omits ADD:

```
proc fcmp outlib=work.funcs.sort;
   subroutine reorder(arr[*] $);
      outargs arr;
      length key $20 num cnt 8;
      declare hash h(ordered: 'ascending');
      declare hiter iter('h');
      rc=h.defineKey('key');
      rc=h.defineData('key','cnt');
      rc=h.defineDone();
      do num=1 to dim(arr);
         key=arr[num];
         * increment/add a key;
         if h.find()=0 then cnt+1;
         else cnt=1;
         rc=h.replace();
         end;
      * build ordered array;
      num=1;
      do while(iter.next()=0);
         do tot=1 to cnt;
            arr[num]=key;
            num+1;
            end;
         end;
      endsub;
quit;
```

The REORDER subroutine overcomes the 800-element arbitrary restriction of SORTC, and prints the first 80 words (alphabetically) occurring in the U.S. Constitution:

```
data const_wide_sorted;
```

```
   set const_wide (drop=_name_);
   array words[*] $ word:;
   call reorder(of words);
   put word1-word80;
run;
```

The log shows the sorted values of Word1 through Word80:

```
a a a a a a a a a a a a a a a a a a a a a a a a a a a a a a a a a a a a a a a
a a a a a a a a ability absence absent absolutely accept acceptance according according
according according accordingly account act act act acts acts actual actual actually
adding adhering adjourn adjourn adjourn adjournment adjournment adjournment
NOTE: There were 1 observations read from the data set WORK.CONST_WIDE.
NOTE: The data set WORK.CONST_WIDE_SORTED has 1 observations and 4422 variables.
```

It should be noted again that the 800-element limitation of arrays occurs only inside the FCMP procedure; both SORTC and SORTN function brilliantly within the DATA step, and can fully leverage the OF operator.

PERFORMING A HORIZONTAL HASH SORT OF A CHARACTER VARIABLE

The previous example demonstrates how to perform a horizontal sort on a series of variables inside a user-defined function or subroutine. And because multi-element arguments must be passed inside an array, this requires an array declaration inside the calling program (i.e., the DATA step), after which the array must be passed by reference (using the OUTARGS statement) to the function or subroutine, after which an ordered hash object must be iterated to reorder the array elements.

In other cases, however, you may want to sort the elements (i.e., words) within a single character variable, and the hash object also facilitates this functionality. For example, consider the following DATA step, which attempts to leverage SORTC to reorder the words in the String variable:

```
data _null_;
   length some_words $100;
   some_words='curry petra cheap friends inside of abdullah aqib bought for and';
   call sortc(some_words);
   put some_words=;
run;
```

This attempt fails to produce the desired result because SORTC is intended to reorder separate variables or values—not separate words within a single variable:

```
some_words=curry petra cheap friends inside of abdullah aqib bought for and
```

However, a user-defined subroutine can be designed that produces the desired functionality. Again, because of the limitations demonstrated previously in both SAS arrays and the OF operator within the FCMP procedure, a hash object must deliver this functionality. An array-based solution could be engineered to sort strings that contain 800 or fewer words; however, as the number of words surpasses 800, array-based solutions will fail.

The REORDER_WORDS user-defined subroutine now sorts the individual words (i.e., space-delimited alphanumeric characters) within a single character variable:

```
proc fcmp outlib=work.funcs.sort;
   subroutine reorder_words(string $);
      outargs string;
      length key $20 num cnt 8;
      declare hash h(ordered: 'ascending');
      declare hiter iter('h');
      rc=h.defineKey('key');
      rc=h.defineData('key','cnt');
      rc=h.defineDone();
      do num=1 to countw(string);
         key=scan(string,num,'','S');
         * increment/add a key;
         if h.find()=0 then cnt+1;
         else cnt=1;
```

```
        rc=h.replace();
        end;
    * rebuild ordered character variable;
    num=1;
    string='';
    do while(iter.next()=0);
        do tot=1 to cnt;
            string=catx(' ',string,key);
            num+1;
            end;
        end;
    endsub;
quit;
```

Note the aberration that the SCAN function must include empty quotes to represent optional arguments that are omitted; this differs from how SCAN, COMPRESS, and other built-in functions are called inside the DATA step:

```
key=scan(string,num,'','S');
```

As before, the hash object reorders the individual words, and repeated words increment the Cnt variable so they can be rewritten multiple times in the String variable. The following DATA step successfully reorders the String variable:

```
data _null_;
   length some_words $100;
   some_words='curry petra cheap friends inside of abdullah aqib bought for and';
   call reorder_words(some_words);
   put some_words=;
run;
```

The log demonstrates that we have aided Abdullah and Aqib in their quest for curry at the unparalleled archaeological site of Petra!:

```
some_words=abdullah and aqib bought cheap curry for friends inside of petra
```

FURTHER FAILURES OF CHARACTER AND NUMERIC ARRAYS IN FCMP

It must be noted that the failure of the OF operator inside the FCMP procedure affects not only SORTC and SORTN but all built-in functions and subroutines, as well as both character and numeric arrays. For example, consider the need to concatenate variables together using the CATX function, which can also leverage the OF operator to specify an array to be concatenated. The following CATX function performs admirably inside the DATA step, and prints the entire text of the U.S. Constitution as a single character variable (Long_var):

```
data _null_;
   set const_wide (drop=_name_);
   length long_var $30000;
   array words[*] $ word:;
   long_var=catx(' ',of words[*]);
   put long_var;
run;
```

However, attempts to replicate this functionality inside the FCMP procedure fail on multiple accounts—identical to the failures demonstrated previously. For example, the OF operator is incompatible with arrays passed from the DATA step:

```
proc fcmp outlib=work.funcs.test;
   function squish_together(arr[*] $) $;
      length so_long $30000;
      so_long=catx(' ',of arr[*]);
      return(so_long);
      endfunc;
quit;
```

The function does not compile, and the log demonstrates this failure:

```
1807  proc fcmp outlib=work.funcs.test;
1808     function squish_together(arr[*] $) $;
1809        length so_long $30000;
1810        so_long=catx(' ',of arr[*]);
ERROR: The OF operator isn't allowed on an ARRAY with a dynamic size.
1811        return(so_long);
1812        endfunc;
1813  quit;

NOTE: The SAS System stopped processing this step because of errors.
```

A second attempted function does compile because the OF operator specifies an array (New_array) that is declared inside the function using the ARRAY statement:

```
proc fcmp outlib=work.funcs.test;
   function squish_together(arr[*] $) $;
      length so_long $30000;
      * copy array argument into new array;
      array new_array[5000] $20;
      do cnt=1 to min(dim(arr),dim(new_array));
         new_array[cnt]=arr[cnt];
         end;
      * attempt to concatenate new array;
      so_long=catx(' ',of new_array[*]);
      return(so_long);
      endfunc;
quit;
```

However, any attempt to pass an array that includes more than 800 elements (i.e., variables) again fails due to the arbitrary limitations of the OF operator inside the FCMP procedure:

```
data _null_;
   set const_wide (drop=_name_);
   length long_var $30000;
   array words[*] $ word:;
   long_var=squish_together(words);
   put long_var;
run;
```

The log shows the failure of OF:

```
1857  data _null_;
1858     set const_wide (drop=_name_);
1859     length long_var $30000;
1860     array words[*] $ word:;
1861     long_var=squish_together(words);
1862     put long_var;
1863  run;

ERROR: Function or Subroutine CATX expects 5001 arguments, this exceeds the maximum
number of arguments allowed, 800.
NOTE: The SAS System stopped processing this step because of errors.
```

One final set of examples demonstrates the full functionality of the MEANS built-in function, which also can leverage the OF operator to specify an array within the DATA step:

```
data nums (drop=cnt);
   length cnt my_mean 8;
   array num[1000];
   call streaminit(123);
   do cnt=1 to 1000;
      num[cnt]=int(100*rand('uniform'));
      end;
   my_mean=mean(of num[*]);
   put my_mean=;
```

```
run;
```

At least within the DATA step, built-in functions and subroutines can typically operate on near-limitless arrays, and the usual limiting factor is one of system resources, such as this failed attempt to expand the Num array to 100 million variables:

```
1999  data nums (drop=cnt);
2000     length cnt my_mean 8;
2001     array num[100000000];
FATAL: Insufficient memory to execute DATA step program. Aborted during the COMPILATION
phase.
ERROR: The SAS System stopped processing this step because of insufficient memory.
NOTE: DATA statement used (Total process time):
      real time            1:01.20
      cpu time             57.34 seconds

2002     call streaminit(123);
2003     do cnt=1 to 100000000;
2004        num[cnt]=int(100*rand('uniform'));
2005        end;
2006     my_mean=mean(of num[*]);
2007     put my_mean=;
2008  run;
```

However, the same pernicious limitations that affect character arrays (and the OF operator) inside the FCMP procedure also affect numeric arrays.

Here, an attempt to call the built-in MEAN function on the Arr_num numeric array fails:

```
proc fcmp outlib=work.funcs.math;
   function avg(arr_num[*]);
      length avg 8;
      avg=mean(of arr_num[*]);
      return(avg);
      endfunc;
quit;
```

The function does not compile, and the log demonstrates the failure:

```
1921  proc fcmp outlib=work.funcs.math;
1922     function avg(arr_num[*]);
1923        length avg 8;
1924        avg=mean(of arr_num[*]);
ERROR: The OF operator isn't allowed on an ARRAY with a dynamic size.
1925        return(avg);
1926        endfunc;
1927  quit;

NOTE: The SAS System stopped processing this step because of errors.
```

Similarly, an attempt to copy the array argument into a new array also fails, as demonstrated previously with character arrays:

```
proc fcmp outlib=work.funcs.math;
   function avg(arr_num[*]);
      length cnt avg 8;
      array new_array[5000];
      do cnt=1 to min(dim(arr_num),dim(new_array));
         new_array[cnt]=arr_num[cnt];
         end;
      avg=mean(of new_array[*]);
      return(avg);
      endfunc;
quit;
```

The function compiles; however, the attempt to call it from the DATA step fails:

```
data nums (drop=cnt);
   length cnt my_mean 8;
   array num[1000];
   call streaminit(123);
   do cnt=1 to 1000;
      num[cnt]=int(100*rand('uniform'));
      end;
   my_mean=avg(num);
   put my_mean=;
run;
```

The log demonstrates this failure—again because of the arbitrary 800-element limitation imposed on arrays inside the FCMP procedure:

```
2009   data nums (drop=cnt);
2010      length cnt my_mean 8;
2011      array num[1000];
2012      call streaminit(123);
2013      do cnt=1 to 1000;
2014         num[cnt]=int(100*rand('uniform'));
2015         end;
2016      my_mean=avg(num);
2017      put my_mean=;
2018   run;
```

ERROR: Function or Subroutine MEAN expects 5000 arguments, this exceeds the maximum
number of arguments allowed, 800.
NOTE: The SAS System stopped processing this step because of errors.

No workarounds are demonstrated in this text for the failures of CATX and MEAN within the FCMP procedure; however, solutions likely could be engineered for these and other built-in functions that fail to process arrays inside the FCMP procedure.

CONCLUSION

This text introduced three failure patterns of arrays that occur inside user-defined functions and subroutines, and which affect both character and numeric arrays called by built-in functions and subroutines. Two failures occur because the OF operator is incompatible with arrays passed as arguments, and the third failure occurs only when arrays exceed 800 elements (i.e., variables) in length. Collectively, these failures prevent longer arrays from being processed effectively as arrays. As a workaround, the hash object (and a bit of clever design) overcomes some limitations. For example, the inability to call the SORTC or SORTN subroutines on longer arrays is overcome by instantiating an ordered hash object, populating it with array values, and by returning the ordered list of values to the calling program.

CONTACT INFORMATION

Your comments and questions are valued and encouraged. Contact the author at:

> Name: Troy Martin Hughes
> E-mail: troymartinhughes@gmail.com

SAS and all other SAS Institute Inc. product or service names are registered trademarks of
SAS Institute Inc. in the USA and other countries. ® indicates USA registration.

Other brand and product names are trademarks of their respective companies.

APPENDIX A. THE U.S. CONSTITUTION

The text of the U.S. Constitution can be saved to the Constitution.txt text file:

We the People of the United States, in Order to form a more perfect Union, establish Justice, insure domestic Tranquility, provide for the common defence, promote the general Welfare, and secure the Blessings of Liberty to ourselves and our Posterity, do ordain and establish this Constitution for the United States of America. Article. I. Section. 1. All legislative Powers herein granted shall be vested in a Congress of the United States, which shall consist of a Senate and House of Representatives. Section. 2. The House of Representatives shall be composed of Members chosen every second Year by the People of the several States, and the Electors in each State shall have the Qualifications requisite for Electors of the most numerous Branch of the State Legislature. No Person shall be a Representative who shall not have attained to the Age of twenty five Years, and been seven Years a Citizen of the United States, and who shall not, when elected, be an Inhabitant of that State in which he shall be chosen. Representatives and direct Taxes shall be apportioned among the several States which may be included within this Union, according to their respective Numbers, which shall be determined by adding to the whole Number of free Persons, including those bound to Service for a Term of Years, and excluding Indians not taxed, three fifths of all other Persons. The actual Enumeration shall be made within three Years after the first Meeting of the Congress of the United States, and within every subsequent Term of ten Years, in such Manner as they shall by Law direct. The Number of Representatives shall not exceed one for every thirty Thousand, but each State shall have at Least one Representative; and until such enumeration shall be made, the State of New Hampshire shall be entitled to chuse three, Massachusetts eight, Rhode-Island and Providence Plantations one, Connecticut five, New-York six, New Jersey four, Pennsylvania eight, Delaware one, Maryland six, Virginia ten, North Carolina five, South Carolina five, and Georgia three. When vacancies happen in the Representation from any State, the Executive Authority thereof shall issue Writs of Election to fill such Vacancies. The House of Representatives shall chuse their Speaker and other Officers; and shall have the sole Power of Impeachment. Section. 3. The Senate of the United States shall be composed of two Senators from each State, chosen by the Legislature thereof, for six Years; and each Senator shall have one Vote. Immediately after they shall be assembled in Consequence of the first Election, they shall be divided as equally as may be into three Classes. The Seats of the Senators of the first Class shall be vacated at the Expiration of the second Year, of the second Class at the Expiration of the fourth Year, and of the third Class at the Expiration of the sixth Year, so that one third may be chosen every second Year; and if Vacancies happen by Resignation, or otherwise, during the Recess of the Legislature of any State, the Executive thereof may make temporary Appointments until the next Meeting of the Legislature, which shall then fill such Vacancies. No Person shall be a Senator who shall not have attained to the Age of thirty Years, and been nine Years a Citizen of the United States, and who shall not, when elected, be an Inhabitant of that State for which he shall be chosen. The Vice President of the United States shall be President of the Senate, but shall have no Vote, unless they be equally divided. The Senate shall chuse their other Officers, and also a President pro tempore, in the Absence of the Vice President, or when he shall exercise the Office of President of the United States. The Senate shall have the sole Power to try all Impeachments. When sitting for that Purpose, they shall be on Oath or Affirmation. When the President of the United States is tried, the Chief Justice shall preside: And no Person shall be convicted without the Concurrence of two thirds of the Members present. Judgment in Cases of Impeachment shall not extend further than to removal from Office, and disqualification to hold and enjoy any Office of honor, Trust or Profit under the United States: but the Party convicted shall nevertheless be liable and subject to Indictment, Trial, Judgment and Punishment, according to Law. Section. 4. The Times, Places and Manner of holding Elections for Senators and Representatives, shall be prescribed in each State by the Legislature thereof; but the Congress may at any time by Law make or alter such Regulations, except as to the Places of chusing Senators. The Congress shall assemble at least once in every Year, and such Meeting shall be on the first Monday in December, unless they shall by Law appoint a different Day. Section. 5. Each House shall be the Judge of the Elections, Returns and Qualifications of its own Members, and a Majority of each shall constitute a Quorum to do Business; but a smaller Number may adjourn from day to day, and may be authorized to compel the Attendance of absent Members, in such Manner, and under such Penalties as each House may provide. Each House may determine the Rules of its Proceedings, punish its Members for disorderly Behaviour, and, with the Concurrence of two thirds, expel a Member. Each House shall keep a Journal of its Proceedings, and from time to time publish the same, excepting such Parts as may

in their Judgment require Secrecy; and the Yeas and Nays of the Members of either House on any question shall, at the Desire of one fifth of those Present, be entered on the Journal. Neither House, during the Session of Congress, shall, without the Consent of the other, adjourn for more than three days, nor to any other Place than that in which the two Houses shall be sitting. Section. 6. The Senators and Representatives shall receive a Compensation for their Services, to be ascertained by Law, and paid out of the Treasury of the United States. They shall in all Cases, except Treason, Felony and Breach of the Peace, be privileged from Arrest during their Attendance at the Session of their respective Houses, and in going to and returning from the same; and for any Speech or Debate in either House, they shall not be questioned in any other Place. No Senator or Representative shall, during the Time for which he was elected, be appointed to any civil Office under the Authority of the United States, which shall have been created, or the Emoluments whereof shall have been encreased during such time; and no Person holding any Office under the United States, shall be a Member of either House during his Continuance in Office. Section. 7. All Bills for raising Revenue shall originate in the House of Representatives; but the Senate may propose or concur with Amendments as on other Bills. Every Bill which shall have passed the House of Representatives and the Senate, shall, before it become a Law, be presented to the President of the United States; If he approve he shall sign it, but if not he shall return it, with his Objections to that House in which it shall have originated, who shall enter the Objections at large on their Journal, and proceed to reconsider it. If after such Reconsideration two thirds of that House shall agree to pass the Bill, it shall be sent, together with the Objections, to the other House, by which it shall likewise be reconsidered, and if approved by two thirds of that House, it shall become a Law. But in all such Cases the Votes of both Houses shall be determined by yeas and Nays, and the Names of the Persons voting for and against the Bill shall be entered on the Journal of each House respectively. If any Bill shall not be returned by the President within ten Days (Sundays excepted) after it shall have been presented to him, the Same shall be a Law, in like Manner as if he had signed it, unless the Congress by their Adjournment prevent its Return, in which Case it shall not be a Law. Every Order, Resolution, or Vote to which the Concurrence of the Senate and House of Representatives may be necessary (except on a question of Adjournment) shall be presented to the President of the United States; and before the Same shall take Effect, shall be approved by him, or being disapproved by him, shall be repassed by two thirds of the Senate and House of Representatives, according to the Rules and Limitations prescribed in the Case of a Bill. Section. 8. The Congress shall have Power To lay and collect Taxes, Duties, Imposts and Excises, to pay the Debts and provide for the common Defence and general Welfare of the United States; but all Duties, Imposts and Excises shall be uniform throughout the United States; To borrow Money on the credit of the United States; To regulate Commerce with foreign Nations, and among the several States, and with the Indian Tribes; To establish an uniform Rule of Naturalization, and uniform Laws on the subject of Bankruptcies throughout the United States; To coin Money, regulate the Value thereof, and of foreign Coin, and fix the Standard of Weights and Measures; To provide for the Punishment of counterfeiting the Securities and current Coin of the United States; To establish Post Offices and post Roads; To promote the Progress of Science and useful Arts, by securing for limited Times to Authors and Inventors the exclusive Right to their respective Writings and Discoveries; To constitute Tribunals inferior to the supreme Court; To define and punish Piracies and Felonies committed on the high Seas, and Offences against the Law of Nations; To declare War, grant Letters of Marque and Reprisal, and make Rules concerning Captures on Land and Water; To raise and support Armies, but no Appropriation of Money to that Use shall be for a longer Term than two Years; To provide and maintain a Navy; To make Rules for the Government and Regulation of the land and naval Forces; To provide for calling forth the Militia to execute the Laws of the Union, suppress Insurrections and repel Invasions; To provide for organizing, arming, and disciplining, the Militia, and for governing such Part of them as may be employed in the Service of the United States, reserving to the States respectively, the Appointment of the Officers, and the Authority of training the Militia according to the discipline prescribed by Congress; To exercise exclusive Legislation in all Cases whatsoever, over such District (not exceeding ten Miles square) as may, by Cession of particular States, and the Acceptance of Congress, become the Seat of the Government of the United States, and to exercise like Authority over all Places purchased by the Consent of the Legislature of the State in which the Same shall be, for the Erection of Forts, Magazines, Arsenals, dock-Yards, and other needful Buildings;—And To make all Laws which shall be necessary and proper for carrying into Execution the foregoing Powers, and all other Powers vested by this Constitution in the Government of the United States, or in any Department or Officer thereof. Section. 9. The Migration or Importation of such Persons as any of the States now existing shall think proper to

admit, shall not be prohibited by the Congress prior to the Year one thousand eight hundred and eight, but a Tax or duty may be imposed on such Importation, not exceeding ten dollars for each Person. The Privilege of the Writ of Habeas Corpus shall not be suspended, unless when in Cases of Rebellion or Invasion the public Safety may require it. No Bill of Attainder or ex post facto Law shall be passed. No Capitation, or other direct, Tax shall be laid, unless in Proportion to the Census or enumeration herein before directed to be taken. No Tax or Duty shall be laid on Articles exported from any State. No Preference shall be given by any Regulation of Commerce or Revenue to the Ports of one State over those of another: nor shall Vessels bound to, or from, one State, be obliged to enter, clear, or pay Duties in another. No Money shall be drawn from the Treasury, but in Consequence of Appropriations made by Law; and a regular Statement and Account of the Receipts and Expenditures of all public Money shall be published from time to time. No Title of Nobility shall be granted by the United States: And no Person holding any Office of Profit or Trust under them, shall, without the Consent of the Congress, accept of any present, Emolument, Office, or Title, of any kind whatever, from any King, Prince, or foreign State. Section. 10. No State shall enter into any Treaty, Alliance, or Confederation; grant Letters of Marque and Reprisal; coin Money; emit Bills of Credit; make any Thing but gold and silver Coin a Tender in Payment of Debts; pass any Bill of Attainder, ex post facto Law, or Law impairing the Obligation of Contracts, or grant any Title of Nobility. No State shall, without the Consent of the Congress, lay any Imposts or Duties on Imports or Exports, except what may be absolutely necessary for executing it's inspection Laws: and the net Produce of all Duties and Imposts, laid by any State on Imports or Exports, shall be for the Use of the Treasury of the United States; and all such Laws shall be subject to the Revision and Controul of the Congress. No State shall, without the Consent of Congress, lay any Duty of Tonnage, keep Troops, or Ships of War in time of Peace, enter into any Agreement or Compact with another State, or with a foreign Power, or engage in War, unless actually invaded, or in such imminent Danger as will not admit of delay. Article. II. Section. 1. The executive Power shall be vested in a President of the United States of America. He shall hold his Office during the Term of four Years, and, together with the Vice President, chosen for the same Term, be elected, as follows Each State shall appoint, in such Manner as the Legislature thereof may direct, a Number of Electors, equal to the whole Number of Senators and Representatives to which the State may be entitled in the Congress: but no Senator or Representative, or Person holding an Office of Trust or Profit under the United States, shall be appointed an Elector. The Electors shall meet in their respective States, and vote by Ballot for two Persons, of whom one at least shall not be an Inhabitant of the same State with themselves. And they shall make a List of all the Persons voted for, and of the Number of Votes for each; which List they shall sign and certify, and transmit sealed to the Seat of the Government of the United States, directed to the President of the Senate. The President of the Senate shall, in the Presence of the Senate and House of Representatives, open all the Certificates, and the Votes shall then be counted. The Person having the greatest Number of Votes shall be the President, if such Number be a Majority of the whole Number of Electors appointed; and if there be more than one who have such Majority, and have an equal Number of Votes, then the House of Representatives shall immediately chuse by Ballot one of them for President; and if no Person have a Majority, then from the five highest on the List the said House shall in like Manner chuse the President. But in chusing the President, the Votes shall be taken by States, the Representation from each State having one Vote; A quorum for this Purpose shall consist of a Member or Members from two thirds of the States, and a Majority of all the States shall be necessary to a Choice. In every Case, after the Choice of the President, the Person having the greatest Number of Votes of the Electors shall be the Vice President. But if there should remain two or more who have equal Votes, the Senate shall chuse from them by Ballot the Vice President. The Congress may determine the Time of chusing the Electors, and the Day on which they shall give their Votes; which Day shall be the same throughout the United States. No Person except a natural born Citizen, or a Citizen of the United States, at the time of the Adoption of this Constitution, shall be eligible to the Office of President; neither shall any Person be eligible to that Office who shall not have attained to the Age of thirty five Years, and been fourteen Years a Resident within the United States. In Case of the Removal of the President from Office, or of his Death, Resignation, or Inability to discharge the Powers and Duties of the said Office, the Same shall devolve on the Vice President, and the Congress may by Law provide for the Case of Removal, Death, Resignation or Inability, both of the President and Vice President, declaring what Officer shall then act as President, and such Officer shall act accordingly, until the Disability be removed, or a President shall be elected. The President shall, at stated Times, receive for his Services, a Compensation, which shall neither be encreased nor diminished during the Period for which he shall have been elected, and he shall not

receive within that Period any other Emolument from the United States, or any of them. Before he enter on the Execution of his Office, he shall take the following Oath or Affirmation:—"I do solemnly swear (or affirm) that I will faithfully execute the Office of President of the United States, and will to the best of my Ability, preserve, protect and defend the Constitution of the United States." Section. 2. The President shall be Commander in Chief of the Army and Navy of the United States, and of the Militia of the several States, when called into the actual Service of the United States; he may require the Opinion, in writing, of the principal Officer in each of the executive Departments, upon any Subject relating to the Duties of their respective Offices, and he shall have Power to grant Reprieves and Pardons for Offences against the United States, except in Cases of Impeachment. He shall have Power, by and with the Advice and Consent of the Senate, to make Treaties, provided two thirds of the Senators present concur; and he shall nominate, and by and with the Advice and Consent of the Senate, shall appoint Ambassadors, other public Ministers and Consuls, Judges of the supreme Court, and all other Officers of the United States, whose Appointments are not herein otherwise provided for, and which shall be established by Law: but the Congress may by Law vest the Appointment of such inferior Officers, as they think proper, in the President alone, in the Courts of Law, or in the Heads of Departments. The President shall have Power to fill up all Vacancies that may happen during the Recess of the Senate, by granting Commissions which shall expire at the End of their next Session. Section. 3. He shall from time to time give to the Congress Information of the State of the Union, and recommend to their Consideration such Measures as he shall judge necessary and expedient; he may, on extraordinary Occasions, convene both Houses, or either of them, and in Case of Disagreement between them, with Respect to the Time of Adjournment, he may adjourn them to such Time as he shall think proper; he shall receive Ambassadors and other public Ministers; he shall take Care that the Laws be faithfully executed, and shall Commission all the Officers of the United States. Section. 4. The President, Vice President and all civil Officers of the United States, shall be removed from Office on Impeachment for, and Conviction of, Treason, Bribery, or other high Crimes and Misdemeanors. Article. III. Section. 1. The judicial Power of the United States, shall be vested in one supreme Court, and in such inferior Courts as the Congress may from time to time ordain and establish. The Judges, both of the supreme and inferior Courts, shall hold their Offices during good Behaviour, and shall, at stated Times, receive for their Services, a Compensation, which shall not be diminished during their Continuance in Office. Section. 2. The judicial Power shall extend to all Cases, in Law and Equity, arising under this Constitution, the Laws of the United States, and Treaties made, or which shall be made, under their Authority;—to all Cases affecting Ambassadors, other public Ministers and Consuls;—to all Cases of admiralty and maritime Jurisdiction;—to Controversies to which the United States shall be a Party;—to Controversies between two or more States;— between a State and Citizens of another State,—between Citizens of different States,—between Citizens of the same State claiming Lands under Grants of different States, and between a State, or the Citizens thereof, and foreign States, Citizens or Subjects. In all Cases affecting Ambassadors, other public Ministers and Consuls, and those in which a State shall be Party, the supreme Court shall have original Jurisdiction. In all the other Cases before mentioned, the supreme Court shall have appellate Jurisdiction, both as to Law and Fact, with such Exceptions, and under such Regulations as the Congress shall make. The Trial of all Crimes, except in Cases of Impeachment, shall be by Jury; and such Trial shall be held in the State where the said Crimes shall have been committed; but when not committed within any State, the Trial shall be at such Place or Places as the Congress may by Law have directed. Section. 3. Treason against the United States, shall consist only in levying War against them, or in adhering to their Enemies, giving them Aid and Comfort. No Person shall be convicted of Treason unless on the Testimony of two Witnesses to the same overt Act, or on Confession in open Court. The Congress shall have Power to declare the Punishment of Treason, but no Attainder of Treason shall work Corruption of Blood, or Forfeiture except during the Life of the Person attainted. Article. IV. Section. 1. Full Faith and Credit shall be given in each State to the public Acts, Records, and judicial Proceedings of every other State. And the Congress may by general Laws prescribe the Manner in which such Acts, Records and Proceedings shall be proved, and the Effect thereof. Section. 2. The Citizens of each State shall be entitled to all Privileges and Immunities of Citizens in the several States. A Person charged in any State with Treason, Felony, or other Crime, who shall flee from Justice, and be found in another State, shall on Demand of the executive Authority of the State from which he fled, be delivered up, to be removed to the State having Jurisdiction of the Crime. No Person held to Service or Labour in one State, under the Laws thereof, escaping into another, shall, in Consequence of any Law or Regulation therein, be discharged from such Service or Labour, but shall be delivered up on Claim of the Party to whom such Service or Labour may be due. Section. 3. New

States may be admitted by the Congress into this Union; but no new State shall be formed or erected within the Jurisdiction of any other State; nor any State be formed by the Junction of two or more States, or Parts of States, without the Consent of the Legislatures of the States concerned as well as of the Congress. The Congress shall have Power to dispose of and make all needful Rules and Regulations respecting the Territory or other Property belonging to the United States; and nothing in this Constitution shall be so construed as to Prejudice any Claims of the United States, or of any particular State. Section. 4. The United States shall guarantee to every State in this Union a Republican Form of Government, and shall protect each of them against Invasion; and on Application of the Legislature, or of the Executive (when the Legislature cannot be convened) against domestic Violence. Article. V. The Congress, whenever two thirds of both Houses shall deem it necessary, shall propose Amendments to this Constitution, or, on the Application of the Legislatures of two thirds of the several States, shall call a Convention for proposing Amendments, which, in either Case, shall be valid to all Intents and Purposes, as Part of this Constitution, when ratified by the Legislatures of three fourths of the several States, or by Conventions in three fourths thereof, as the one or the other Mode of Ratification may be proposed by the Congress; Provided that no Amendment which may be made prior to the Year One thousand eight hundred and eight shall in any Manner affect the first and fourth Clauses in the Ninth Section of the first Article; and that no State, without its Consent, shall be deprived of its equal Suffrage in the Senate. Article. VI. All Debts contracted and Engagements entered into, before the Adoption of this Constitution, shall be as valid against the United States under this Constitution, as under the Confederation. This Constitution, and the Laws of the United States which shall be made in Pursuance thereof; and all Treaties made, or which shall be made, under the Authority of the United States, shall be the supreme Law of the Land; and the Judges in every State shall be bound thereby, any Thing in the Constitution or Laws of any State to the Contrary notwithstanding. The Senators and Representatives before mentioned, and the Members of the several State Legislatures, and all executive and judicial Officers, both of the United States and of the several States, shall be bound by Oath or Affirmation, to support this Constitution; but no religious Test shall ever be required as a Qualification to any Office or public Trust under the United States. Article. VII. The Ratification of the Conventions of nine States, shall be sufficient for the Establishment of this Constitution between the States so ratifying the Same. The Word, "the," being interlined between the seventh and eighth Lines of the first Page, The Word "Thirty" being partly written on an Erazure in the fifteenth Line of the first Page, The Words "is tried" being interlined between the thirty second and thirty third Lines of the first Page and the Word "the" being interlined between the forty third and forty fourth Lines of the second Page. Attest William Jackson Secretary