The Astronomical Journal, 147:80 (7pp), 2014 April doi:10.1088/0004-6256/147/4/80

AN EXPERT COMPUTER PROGRAM FOR CLASSIFYING STARS ON THE MK SPECTRAL CLASSIFICATION SYSTEM

R. O. Gray[1] and C. J. Corbally[2]
[1] Department of Physics and Astronomy, Appalachian State University, Boone, NC 26808, USA
[2] Vatican Observatory Research Group, Tucson, AZ 85721-0065, USA
Received 2013 September 5; accepted 2014 January 17; published 2014 March 11

ABSTRACT

This paper describes an expert computer program (MKCLASS) designed to classify stellar spectra on the MK Spectral Classification system in a way similar to humans—by direct comparison with the MK classification standards. Like an expert human classifier, the program first comes up with a rough spectral type, and then refines that spectral type by direct comparison with MK standards drawn from a standards library. A number of spectral peculiarities, including barium stars, Ap and Am stars, λ Bootis stars, carbon-rich giants, etc., can be detected and classified by the program. The program also evaluates the quality of the delivered spectral type. The program currently is capable of classifying spectra in the violet–green region in either the rectified or flux-calibrated format, although the accuracy of the flux calibration is not important. We report on tests of MKCLASS on spectra classified by human classifiers; those tests suggest that over the entire HR diagram, MKCLASS will classify in the temperature dimension with a precision of 0.6 spectral subclass, and in the luminosity dimension with a precision of about one half of a luminosity class. These results compare well with human classifiers.

Key words: stars: fundamental parameters – stars: general – stars: peculiar – techniques: spectroscopic

1. INTRODUCTION

MK Spectral Classification, first developed by Morgan et al. (1943), is the primary system astronomers use for classifying stars based on their spectra. Spectral types are useful in the estimation of physical parameters (effective temperature, surface gravity, metal abundances), ages, distances, and reddenings, and have played a critical role in many fields of astrophysics (cf. Gray & Corbally 2009). A second valuable service provided by spectral classification is the identification of peculiar and astrophysically interesting stars. The MK system has been extended beyond the optical realm into the ultraviolet (Walborn et al. 1985, 1995) and the infrared (Hanson et al. 1996; Andrillat et al. 1995; Meyer et al. 1998; Wallace et al. 2000). For a fuller set of references see Gray & Corbally (2009).

Until recently, spectral classification was carried out by eye using glass plates under a dissecting microscope or a spectral comparator. With the adoption of digital detectors, spectral classification moved seamlessly to the computer screen, but the normal mode was still human interaction with individual spectra. With the advent of multifiber spectrographs and large surveys such as the Sloan Digital Survey, SEGUE, Large Sky Area Multi-Object Fiber Spectroscopic Telescope (LAMOST), etc., demand has been growing for automatic classification of spectra. Historically, the development of automatic classification systems has been along two lines: the metric-distance technique (cf. Kurtz 1984; LaSala 1994) and techniques based on Artificial Neural Networks (ANN; von Hippel et al. 1994; Bailer-Jones 1997; Weaver 1994; Gulati et al. 1994; Singh et al. 1998). More recently, fuzzy logic knowledge-based systems have been investigated (Carricajo et al. 2004; Manteiga et al. 2009). Of the three, the metric-distance technique is somewhat more faithful to the practice of MK classification, as it relies on a (weighted) least-squares comparison with the spectra of MK standard stars. The ANN technique, which has been the most popular, requires training with a large set of spectra classified by an expert classifier, and thus has only an indirect connection with the MK standards. As implemented in STARMIND (Manteiga et al. 2009), the fuzzy logic technique references the MK standards through a list of line ratios, line strengths, and band fluxes. The metric distance method and the ANN technique have been quite successful in temperature classification, but luminosity classification has never been entirely satisfactory except in limited regions of the HR diagram. STARMIND, which admittedly is a preliminary attempt to apply fuzzy logic to MK classification, produces only one-dimensional spectral types (no luminosity classes) and those show a rather large dispersion when compared to spectral types by human experts. However, the fuzzy logic system has the advantage of providing a probability value for the final classification, and may have an interesting potential if applied within the context of an expert system such as described in this paper.

There are a number of disadvantages to each of these techniques. The metric-distance method is quite sensitive to the accuracy of the rectification or flux calibration of the spectra involved. Because the ANN technique does not involve a direct comparison with the MK standards, but rather relies on the "cloud" of spectral types by an expert classifier (for instance, Nancy Houk, and the Michigan HD Reclassification project), retraining is necessary if the system is confronted with spectra with a different spectral region or resolution. The fuzzy logic technique (STARMIND) claims to use a human-like approach to classification, but it is not at all clear that the approach used is, indeed, similar to the way that humans actually classify spectra. Perhaps most importantly, no current technique seems capable of performing one of the most important functions of spectral classification—the identification and classification of peculiar and thus astrophysically interesting stars.

Our own involvement with classifying the 3600 spectra of the Nearby Stars (NStars) project (Gray et al. 2003, 2006) and our current project of classifying 22,664 spectra in the *Kepler* field taken with LAMOST (Wang et al. 1996) has convinced us that a simple, adaptable, automatic classification system is necessary. Our most important criterion is that it be faithful to the MK Spectral Classification system, and what this essentially means is that classification must be carried out by direct comparison with the MK standards using techniques and criteria similar to those of human classifiers. One of us (R.O.G.) has undertaken the programming of such an "expert" system. We detail the progress toward that goal in this paper.

2. HOW HUMANS CLASSIFY SPECTRA

The MK Classification System, first devised by Morgan et al. (1943), classifies stars by direct comparison with MK standard stars. The most recent detailed discussion of the MK system is that of Gray & Corbally (2009); that reference also has an appendix with extensive lists of MK standard stars. The practice of classification by humans is straightforward, at least for stars with "normal" spectra. Ideally, a library of standard spectra obtained with the same instrument/telescope combination is available. If such a library is not available, then either the program spectra must be manipulated (e.g., by convolving with an appropriate line-spread function) to match the resolution of the available library, or the library must be transformed by a similar process to match as closely as possible the resolution of the program spectra. Once that is accomplished, individual spectra are classified with the following procedure. A program spectrum is displayed on the computer screen. An experienced classifier can usually at that point assign a rough spectral type (for instance, "this spectrum is a K giant," or even "this spectrum is an early K giant"). Detailed comparison with the MK standards then begins. The program spectrum is displayed on the computer screen with, usually, two standards. The initial effort is to determine an improved temperature type (for instance, "K2") by finding standards (with the same preliminary luminosity class) that bracket the temperature type of the program star. Certain spectral criteria (line ratios, line strengths, etc.) may be employed, but the process is essentially unitary, as the entire spectrum is involved in the comparison between the standard and the program spectrum. Once the temperature type is found, a more precise luminosity type is then determined, again by bracketing with standards, but this time in the luminosity dimension. A more precise temperature type is then found, and so the iteration continues until the best type has been determined. The above procedure involves mental interpolation between the standards, as a standard with the exact spectral type of the program star may not be available.

The human classifier must always consider the possibility that the spectrum is "peculiar" in some sense and thus will not conform closely to the MK standards. A common example is a star—say a G-type dwarf—that is metal-weak. With that possibility in mind, the human classifier first considers criteria that are independent of the metallicity—in a G-type star those are the hydrogen-line strengths and certain Fe I/Cr I ratios (see Gray & Corbally 2009). The classifier then considers the strength and morphology of the metallic-line spectrum, and only if that gives a spectral type that is consistent with the metallicity-independent criteria (implying solar abundances) are the standard temperature-sensitive Fe I/hydrogen ratios used to refine the spectral type. Otherwise, the "metallicity" of the star is estimated by comparison with earlier or later MK standards (Corbally 1987), or with metal-weak and metal-strong standards (Gray 1989). Other types of peculiarities require different procedures. In some cases it is sufficient to determine the best spectral type ignoring the peculiarities and then note those peculiarities in the final spectral type (e.g., B8 V Si). In other cases the peculiarity is so dominant that recourse is made to describing different parts of the spectrum by reference to different spectral standards. An example of this is the spectral type of an Am star, where separate spectral types are provided for the K-line, the hydrogen lines, and the general metallic-line spectrum (e.g., kA1hA7mF2 V). The classification of peculiar stars is important and human classifiers usually take great pains in dealing with them, as those stars are often of great astrophysical interest.

3. MKCLASS—A SPECTRAL CLASSIFICATION PROGRAM

MKCLASS is a computer program written in the "C" language and currently running only on UNIX/Linux platforms that attempts to reproduce as closely as possible the human process in classifying a spectrum. It contains the basic essential elements of an "expert system" (see Jackson 1998) in that it has a *knowledge base* (a library of MK standard spectra) and an *inference engine* consisting of classification modules that carry out the classification, using human-like reasoning, via direct comparison of the unknown and the classification standards. As such, the program conforms to the basic characteristics of a zeroth-order expert system. Below, we discuss the function of that program parallel to our discussion of the human process of classification in the previous section. A flowchart for MKCLASS is illustrated in Figure 1.

MKCLASS is a work in progress; the current discussion is based on version 1.01 which was frozen 2013 August 15.

3.1. Standard Libraries

MKCLASS is currently supplied with two libraries of MK standard spectra. The source of the spectra for these libraries is the GM spectrograph on the 0.8 m reflector of the Dark Sky Observatory situated in northwestern North Carolina. Many of those standard spectra appear in the illustrations in Gray & Corbally (2009). The first standard library, `libr18`, consists of 1.8 Å resolution rectified spectra with a spectral range from 3800–4600 Å; they were obtained with a 1200 g mm^{-1} grating on the GM spectrograph. The second standard library `libnor36` consists of 3.6 Å resolution flux-calibrated and normalized spectra with a spectral range from 3800–5600 Å, and were obtained with a 600 g mm^{-1} grating. While the `libnor36` library spectra were obtained with a narrow-slit spectrograph, efforts were made to ensure that the flux calibrations of the individual spectra are accurate. Fluxes are corrected using Strömgren photometry (for a discussion of this, see Gray et al. 2003) and the corrected flux distributions are dereddened. All fluxes are normalized at a single consistent point (4503 Å). Adding spectral libraries to MKCLASS is straightforward; they are simply described in the file `mkclass.lib` using certain keywords and then they become accessible to MKCLASS.

Currently, MKCLASS is programmed with spectral classification criteria in the violet–green part of the spectrum (3800–5600 Å), but it is anticipated that this spectral range will be extended in later versions. "Essential" spectral criteria are contained in the spectral range 3918–4600 Å for O–K-type stars; if the program stars include M-type stars, that "essential" range extends to 5000 Å. What that means is that MKCLASS currently requires the program spectra to range minimally from 3918–4600 (5000) Å. This corresponds closely to what humans require for accurate spectral classification in the

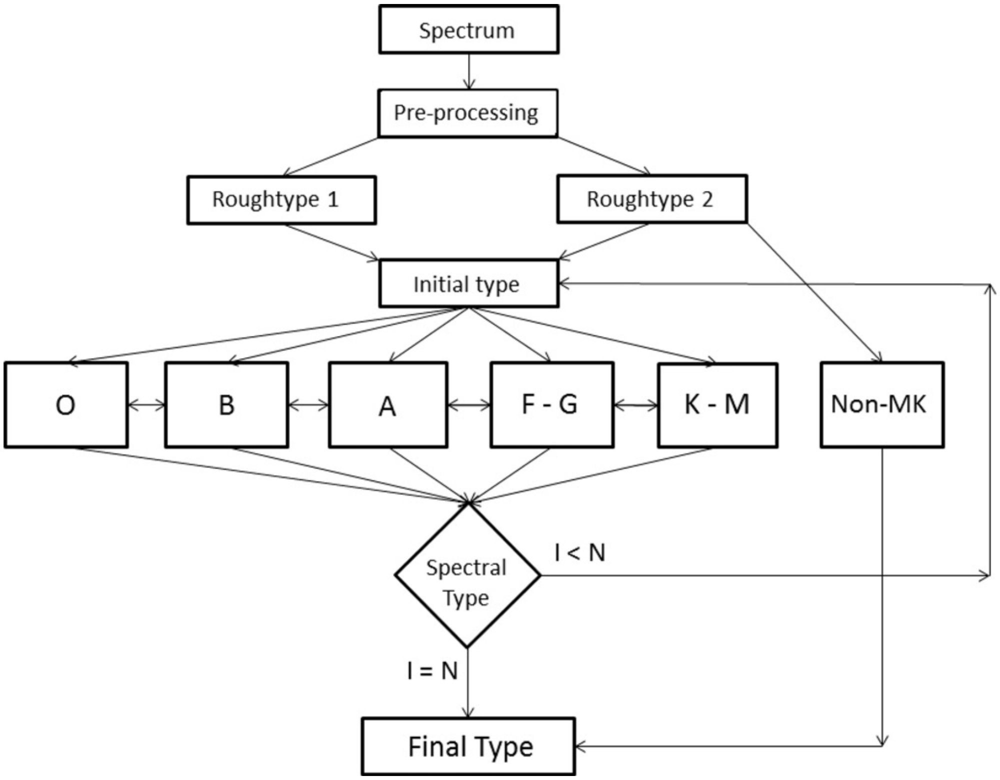

Figure 1. Simplified flow chart for MKCLASS. See the text. MKCLASS can be programmed to iterate through the entire classification process *N* times. The first decision—whether to use ROUGHTYPE 1 or ROUGHTYPE 2—is made once per set of unknowns, according to whether those spectra are rectified or flux calibrated (or uncalibrated), and thus MKCLASS is essentially an *unsupervised* classification program.

blue–violet—classifying A-type stars depends critically on access to the Ca II K-line, and classification of M-type stars depends upon access to at least one TiO band. We are currently working on compiling another standard library, one with a (∼4 Å) resolution but with a much wider spectral range. When that library is available, MKCLASS will be extended to the red part of the spectrum, and criteria will be added that will make the above "essential minimum" spectral range more flexible. The source of this new library is the VATTspec spectrograph on the 1.8 m Vatican Advanced Technology Telescope.

The standard libraries are compiled from MK standard star spectra. MKCLASS expects a fully populated subgrid of standards at pre-specified temperature types and at luminosity classes "V," "III," "Ib," and "Ia." The meaning here of "subgrid" is that each standard library need not span the entire spectral type range (O–M), but should be complete within a limited subrange. That subrange is specified in the file `mkclass.lib`. Of course, MK standards might not be available at all of those grid points, and so interpolation between existing MK standards is employed when building the library. More details can be found in the documentation that accompanies the software. Third-party libraries are strongly encouraged and will be welcomed, just so long as they are based on MK standards.

The MK Spectral Classification system is defined by the MK standards. According to Morgan (1984) "the classification act itself consists of comparisons with the series of standard spectra that define the (spectral-type) boxes with the question: "Is the unknown spectrum (*x*) "like" or "not like" this particular standard spectrum?"" It is this act of comparison with standards that guarantees the MK spectral type is objective and not subjective. MKCLASS conforms to this basic process of classification in that its "knowlege base" consists of the standard spectrum libraries described above. In this sense, MKCLASS deviates somewhat from the normal practice used in the construction of "expert systems," in that the knowledge base is usually expressed in the form of statements in a special-purpose language that encode that knowledge. However, since the essence of MK classification is the *direct* comparison of the unknown with standard spectra, it was felt that expressing the knowledge contained in those standards in terms of "statements" in a special computer language would require an unnecessary deviation from the essential practice of MK classification. Instead, MKCLASS extracts the necessary knowledge from its knowledge base of standard spectra through real-time measurements, and then uses that knowledge in its "inference engine" to decide on a spectral type. How that inference engine works will be described in more detail below.

3.2. Preprocessing of Program Spectra

MKCLASS uses a companion program, MKPRELIM, to carry out some basic preliminary manipulation of the program spectra. MKPRELIM determines the radial velocity of the program spectrum, and transforms the spectrum to the rest frame of the star. MKPRELIM will also normalize (at a specific wavelength) the program spectrum if the spectrum is in the flux or uncalibrated formats. MKPRELIM currently does not convolve the program spectrum with a line-spread function to match it to the resolution of the library spectra; that function needs to be carried out in a separate preprocessor. The software distribution includes the necessary auxiliary programs that can be used to build, in a modular fashion, scripts for pre-processing spectral data sets.

3.3. Derivation of an Initial Rough Spectral Type

Like a human classifier, MKCLASS first determines a rough spectral type. If the spectra to be classified are rectified, a weighted χ^2 difference scheme measured with respect to the library spectra does that job adequately, at least for reasonably "normal" stars. This procedure is referred to as ROUGHTYPE 1, and can be specified on the command line that invokes MKCLASS. However, if the program spectra are flux calibrated or not calibrated at all, then using a χ^2 technique can give spurious answers, and should not be employed. The reason for this is that most classification-resolution spectra are obtained either with a narrow slit or with a fiber (such as the LAMOST spectra), and in both cases the flux calibration, if attempted, is suspect. Reddening can also, of course, affect the fluxes, and this could easily lead a χ^2 technique to give a wildly inaccurate initial type. In addition, if the star is very metal-weak (such as HD 19445) or is very peculiar, then the χ^2 technique fails. In all those cases, it is recommended that ROUGHTYPE 2 be used to determine the initial rough spectral type. The ROUGHTYPE 2 method employs spectral indices to determine a rough spectral type, and can even be used with spectra that are not flux calibrated at all.

ROUGHTYPE 2 begins by distinguishing spectra that are dominated by emission lines from those that are dominated by absorption features. Spectra that cannot be classified on the canonical MK Classification System (white dwarfs, carbon stars, Wolf–Rayet stars, etc.) are then identified in the ROUGHTYPE 2 procedure, and shunted into subroutines which eventually will be developed into full classification routines. Currently, MKCLASS will simply print out a very rough spectral type (DA, DZ, WN, etc.) and then proceed to the next star. The ROUGHTYPE 1 routine does not do that, and so that routine should be used only when it is known that such types are either absent from the spectral database that is being classified or are of little interest to the users. One advantage of ROUGHTYPE 1 is that it is faster than ROUGHTYPE 2.

The decision to use either ROUGHTYPE 1 or ROUGHTYPE 2 to obtain an initial rough spectral type is typically made only once for a given set of unknown spectra. The decision is simple—if the unknown spectra are rectified, use ROUGHTYPE 1, otherwise, use ROUGHTYPE 2. Since that decision is made only once for a given set of unknowns, and does not require expert knowledge, this means that MKCLASS is essentially an "unsupervised" classification program, in that it can classify thousands of spectra in a homogeneous database without any human interaction.

3.4. The Classification Process

Once a rough spectral type has been determined, and non-MK types shunted off, MKCLASS moves onto the process of a detailed comparison of the program spectrum with the standard library. MKCLASS employs five modules to carry out that comparison. These modules constitute the "inference engine" of the expert system. Those five modules are designed to classify O-type stars, B-type stars, A-type stars, F- and G-type stars, and K- and M-type stars. These modules are designed so that they can, if necessary, pass a spectrum to an adjacent module. For example, if the ROUGHTYPE procedure identifies a star as a K-type star, but further classification in the K–M routine indicates it is a G-type star, that module can pass the spectrum to the F–G module. These modules invoke subroutines that directly apply MK spectral classification criteria (in the form of line ratios or spectral indices) appropriate for the spectral type. It is in those subroutines that the direct comparison between the program star and the standard star is made.

Like a human classifier, MKCLASS first considers the possibility that the spectrum in question is peculiar in some sense, and follows steps similar to those outlined in the previous section on human classification. That necessity—to consider spectral peculiarities at every step—is what most clearly distinguishes MKCLASS from other automatic classification methods and also requires a more sophisticated approach to classification than mere least-squares fitting to library spectra. How MKCLASS goes about detecting spectral peculiarities differs from one spectral class to the other. For instance, in the A-type stars, MKCLASS first determines separate temperature types for the Ca II K-line, the hydrogen lines, and then for the general metallic-line spectrum. Only if those are in reasonable agreement does the program assume the spectrum is "normal." If the K-line and metallic-line types are "earlier" than the hydrogen-line type, a metal-weak star, such as a λ Boo star or a horizontal-branch star, is suspected. If the K-line type is earlier and the metallic-line type later than the hydrogen-line type, then the classification is of an Am (metallic-line A-type) star. In both cases the spectrum is passed to specialized routines. MKCLASS also examines certain specific spectral lines, such as Sr II λ4077, Si II λλ4128-30, etc. to detect an Ap star. Details on the spectral criteria employed by MKCLASS across the HR diagram may be found in Gray & Corbally (2009).

It should be emphasized that tables of equivalent widths and/or line ratios as functions of spectral type or luminosity class are not employed by MKCLASS except in the ROUGHTYPE 2 routine. When it is necessary for the program to use a line ratio or spectral index as a classification criterion, that feature is compared directly with the same feature in the library spectra, meaning that that measurement occurs during the classification process. This makes it relatively straightforward to add standard spectral libraries to MKCLASS and thus MKCLASS may be easily adapted to any large spectral survey and also to a wide range of spectral resolutions.

Spectral classification criteria that involve the strengths of spectral features (such as the strength of the Balmer lines, the Ca II K-line, the *G*-band, etc.) are measured as *spectral indices*, i.e., the line flux is measured in a narrow band centered on the feature, and then is ratioed with fluxes in "continuum" bands flanking the feature. The only line ratios used by MKCLASS involve lines with small wavelength separations. The strengths of those lines are, again, measured in narrow bands centered on the lines. This use of spectral indices and closely spaced lines in line ratios frees MKCLASS from the need for accurately rectified or flux-calibrated spectra, in contrast to the metric-distance technique. It also means that MKCLASS does not require the unknown spectrum to be corrected for reddening, and that MKCLASS is tolerant of low signal-to-noise (S/N) spectra (see Section 4.3). Furthermore, these techniques mean that MKCLASS does not, in any context, engage in continuum determination.

MKCLASS does not rely solely on line ratios and strengths to determine the spectral type, but consistent with the "unitary" principle of MK classification, it also employs direct comparisons of several large spectral regions. Which spectral regions are used depends on the spectral type. Those comparisons do not require accurate rectification or flux calibration of the program spectrum, as they are approximately normalized by MKCLASS using a low-order polynomial fit before the comparison occurs.

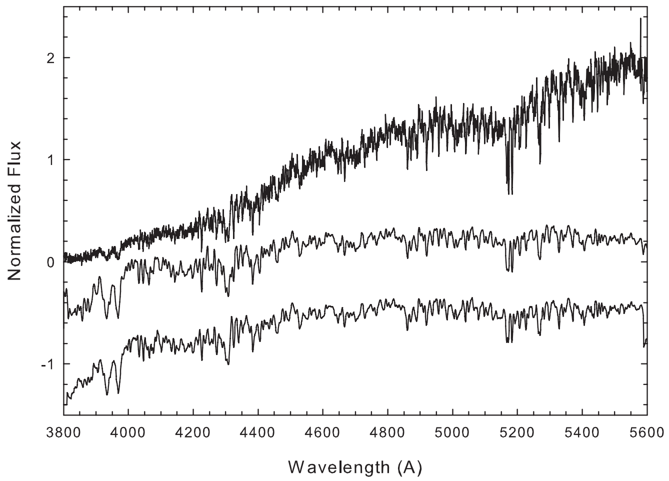

Figure 2. Montage of three spectra showing, from top to bottom, an example of an original uncalibrated LAMOST spectrum, truncated to 3800–5600 Å, and normalized to unity at 4503 Å; the corresponding pre-processed and flux-corrected LAMOST spectrum that is used by MKCLASS to derive the final spectral type; and the interpolated library spectrum identified by MKCLASS as the best match to the LAMOST spectrum. The MKCLASS spectral type for this star is G6 IV.

Like human classifiers, MKCLASS employs interpolation between standard spectra as it converges on a solution. Also like humans, MKCLASS uses an iterative process to reach a precise spectral type. The number of iterations is not fixed, as MKCLASS can perform multiple iterations in passing a given spectrum back and forth between the spectral classification modules. However, once a "final" spectral type is derived, MKCLASS can be programmed to repeat the classification, beginning with that "final" spectral type. The purpose of this is to improve the precision of the classification by beginning the classification process at a better initial point. While MKCLASS is largely insensitive to a poor program-star flux calibration, repeated classifications can be exploited to "improve" that flux calibration, which can help improve the precision of the final type. The way that this is done is to use the best matched (possibly interpolated) library spectrum as a flux template. An option then exists to "adjust" the fluxes of the program star to that flux template before the next classification attempt. This makes possible, as an optional output, a "flux-corrected" version of the program spectrum, which may be of some use in estimating basic stellar parameters (see Figure 2 which illustrates this process for a LAMOST spectrum). It should be emphasized, however, that this flux "improvement" process is not crucial to the function of MKCLASS, and indeed, MKCLASS can accurately classify uncalibrated spectra.

MKCLASS outputs its results to two files. The first file (called the classification file) simply records the final spectral type. If the star is determined to be "normal," MKCLASS also outputs to that file an evaluation of the quality of that classification, based on an overall χ^2 difference between the program spectrum and the best match to that spectrum drawn from the standard library. The quality evaluations are given as "excellent," "very good," "good," "fair," and "poor." Since those evaluations depend on both how well the program spectrum is matched by an (interpolated) standard library spectrum and the S/N of the program spectrum, they are a good indication of how reliable that particular classification is. The second output file is a "log" which contains detailed information on the process MKCLASS employed to arrive at the final spectral type. This log file is useful for debugging MKCLASS, but also gives the user some insight into the nature of a peculiar star. Numerical values for the χ^2 comparison, which is used for the quality evaluation, are also recorded in the log file. For normal stars, MKCLASS also outputs a spectrum, interpolated from the standard library, which is the best match to the program spectrum. It is that "best match" spectrum that forms the basis for the quality evaluation. If instructed to do so, MKCLASS also outputs a "flux-corrected" (see above) version of the program spectrum. MKCLASS can be operated from the command line, making possible batch classification of large numbers of spectra. Since MKCLASS "appends" its output for each star to both the classification file and the log file, the results for large numbers of stars can be output to single classification and log files.

4. TESTS—HOW WELL DOES MKCLASS DO?

4.1. Spectral and Luminosity Classification

To test how well MKCLASS classifies stellar spectra, we have employed three sets of carefully classified stars. The first two sets come from the NStars Project, in particular the northern sample (Gray et al. 2003) which was observed with the Dark Sky Observatory (Appalachian State University) GM spectrograph. Those spectra were obtained with two resolutions, 1.8 Å and 3.6 Å (2 pixel resolutions), using the 1200g mm^{-1} and 600g mm^{-1} gratings, respectively. Stars earlier than G5 were primarily observed with the higher-resolution configuration (spectral range 3800–4600 Å), whereas the fainter, later-type stars were mostly observed with the 600 g mm^{-1} grating (3800–5600 Å). The high-resolution spectra were rectified, and the lower resolution were approximately flux calibrated. Spectral types in the northern sample of the NStars project range from B7–M4. Most of these stars are dwarfs; the most luminous are giants. The spectra have S/N $\gtrapprox$ 100 except for some of the fainter stars, typically M-dwarfs. We have supplemented this sample with a small collection of B-type stars observed at the 1.8 Å resolution. The entire sample contains 960 spectra, 522 at the higher resolution, 438 at the lower.

Because the NStars sample contains only a limited range of luminosity types, we have also tested MKCLASS against a sample of A-, F-, and G-type stars with luminosity types ranging from dwarf to supergiant. The spectral types for this sample (hereafter referred to as the "luminosity sample") were published in Gray et al. (2001). This sample contains a total of 286 spectra, all at the higher resolution.

The spectral standard libraries `libr18` and `libnor36` discussed above are derived from spectra obtained with the same spectrograph/telescope combination as these sets, and these libraries also contain the same standards used by the human classifiers. As a consequence, these tests represent the best results that may be expected from MKCLASS. The reader may question the validity of these comparisons since the human classifiers and the programmers of MKCLASS were the same. We remind the reader of the objectivity of the MK spectral type provided it is obtained through direct comparison with the standards (see the discussion in Section 3.1). Since MKCLASS conforms to this requirement, and carries out its classifications in an unsupervised fashion, we are here comparing two objectively derived sets of MK spectral types, so the comparisons should be valid and representative of the ability of MKCLASS to classify stars.

Application of MKCLASS to these program spectra is straightforward, as no pre-processing is required except for radial-velocity correction, and that is built into MKPRELIM which is called by MKCLASS.

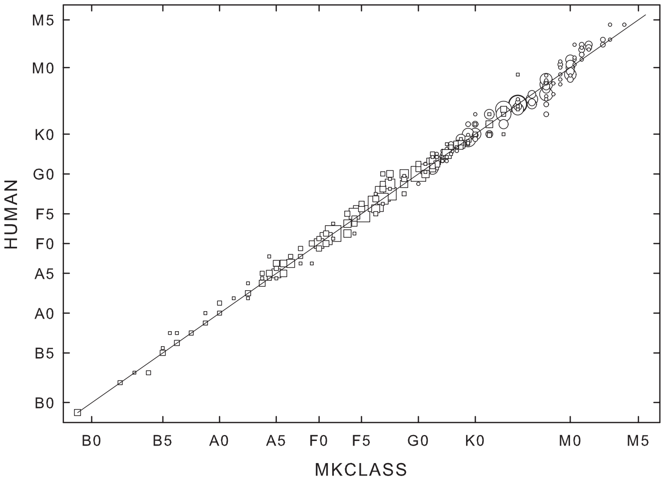

Figure 3. MKCLASS temperature types for the NStars and luminosity samples, plotted against the temperature types of Gray et al. (2001) and Gray et al. (2003). The circles represent spectral types derived from the lower-resolution (3.6 Å) spectra, the squares from the higher-resolution (1.8 Å) spectra. The size (area) of a given symbol is proportional to the number of spectral types represented by that symbol. A total of 1246 spectral types were used to construct this figure.

Figure 3 shows the comparison between the MKCLASS spectral (temperature) types and the published spectral types from both the NStars and the luminosity samples. The size of a symbol in that diagram (specifically the area covered by the symbol) is proportional to the number of stars represented by that symbol. The circles represent spectral types determined from the lower-resolution spectra, the squares from the higher-resolution spectra. The solid line has a slope of unity and an intercept of zero. The standard deviation in the comparison (MKCLASS versus human spectral types) is 0.59 spectral subclass (where the difference, for instance, between F5 and F6 represents one spectral subclass), with a systematic difference between the MKCLASS and human types of only 0.08 ± 0.02 subclass, in the sense that the MKCLASS types are slightly earlier than the human types. The standard deviation and zero-point difference appear to be nearly constant across the entire range of spectral types (O9–M4), and not dependent on the spectral resolution.

The ability of MKCLASS to carry out luminosity classification is best judged from the luminosity sample. The results are illustrated in Figures 4 and 5. Figure 4 shows a histogram of the luminosity class difference between MKCLASS and the human types, where a unit luminosity class difference is the difference between, for instance, a dwarf (V) and a subgiant (IV) classification. The luminosity class difference has a standard deviation of 0.52 luminosity class, with a systematic difference of only 0.02 ± 0.03 luminosity class. Figure 5, which plots the cumulative luminosity-class distributions (in which the ordinate represents the number of stars with luminosity class more luminous than the abscissa) for both the MKCLASS and human types, shows, however, that MKCLASS tends to classify low luminosity stars (dwarfs) on average as slightly more luminous than the human types, and supergiants as slightly less luminous.

These results indicate that MKCLASS is a competent classifier of normal stars.

On a reasonably fast laptop, MKCLASS typically takes a few seconds to carry out a single classification. Most of the computer time is taken in the radial velocity correction and pre-processing.

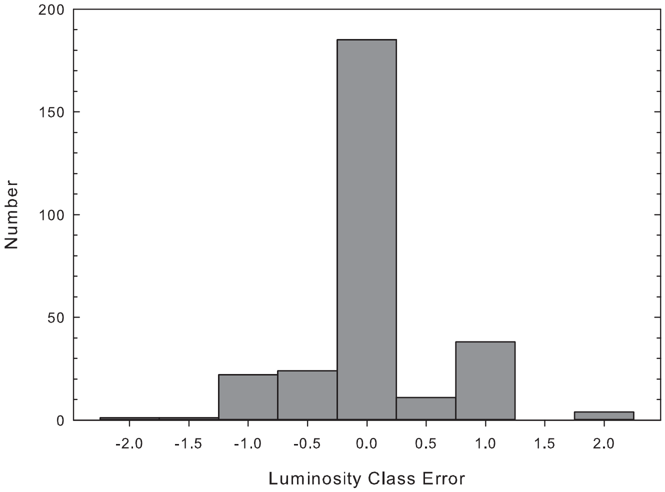

Figure 4. Histogram showing the difference between MKCLASS luminosity classes and the luminosity classes of Gray et al. (2001) for the luminosity sample. A difference of 1 corresponds to a difference between, for instance, a luminosity class of V and one of IV.

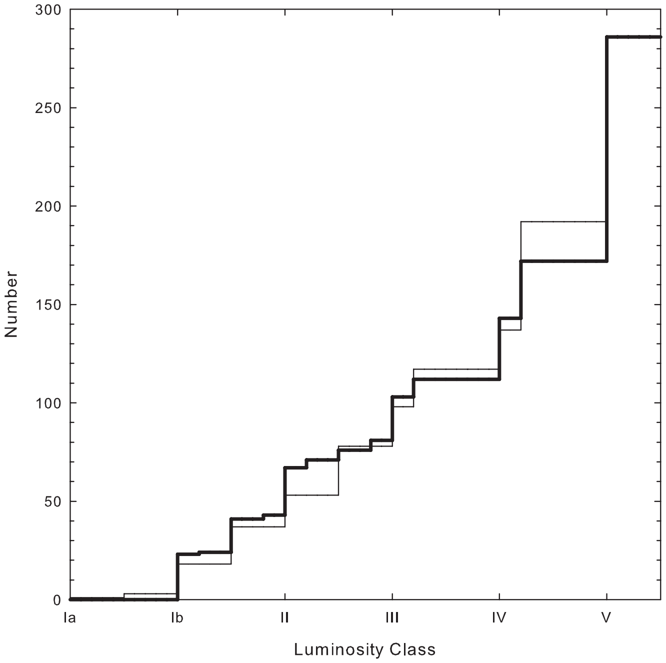

Figure 5. Cumulative luminosity-class distributions for the human spectral types (thick line) and the MKCLASS spectral types (thin line). The cumulative distribution is formed by plotting the number of stars with luminosity classes more luminous than the abscissa. The two distributions are quite similar, but show that MKCLASS tends, on average, to classify dwarfs as slightly more luminous than the human types, and the opposite for the supergiants.

4.2. The Classification of Peculiar Stars

The ability to identify and classify peculiar stars is of equal importance to competence in the classification of normal stars. As described above, MKCLASS is designed to detect many of the more common spectral peculiarities. MKCLASS employs a simplified spectral-type notation for peculiar stars. Table 1 compares MKCLASS spectral types of selected peculiar stars with human types.

4.3. Classification Accuracy and Signal-to-noise

How well does MKCLASS perform for low S/N spectra? To test this, we smoothed the 3800–4620 Å section of the Kurucz

Table 1
Spectral-type Comparison for Peculiar Stars

Star ID	HD	Human Type	MKCLASS
	16460	F2 IV-V SrEuCr:	F1 IV-V SrEu
	26367	F7 V^+ Sr II CH+0.4	F7 V Sr
	78362	kA5hF0mF5 II	kA3hA9mF3
μ Leo	85503	K2 III CN2	K3 III CN2
37 Com	112989	G9 III CH-2	G9 III-IV CH-2
λ Boo	125162	A3 Va kB9mB9 Lam Boo	A3 V mB9.5 Lam Boo
HR 6791	166208	G8 III CN-1 CH-3	G8 III-IV CH-3
ϕ Dra	170000	B8 V Si	B8 V Si
	217143	G9.5 III Ba 2+	G9 III CN1 Ba
HR 8799	218396	F0 V kA5mA5 Lam Boo	F0 V mA4 Lam Boo

Table 2
S/N and the Solar Spectral Type

S/N	Spectral Type	Quality
∞	G2 V	Vgood
300	G2 V	Vgood
100	G2 V	Vgood
50	G2 V	Vgood
20	G2 V	Vgood
10	G3 V	Good
5	F8–G8 V	Fair–poor

et al. (1984) solar flux atlas to the resolution of the `libr18` spectral library (1.8 Å). This gives an extremely high S/N spectrum of the Sun which may be classified with MKCLASS. Not surprisingly, MKCLASS returned a spectral type of G2 V. We then applied successively greater amounts of noise to that spectrum and submitted the resulting spectra to MKCLASS. Table 2 shows the results. For the lowest S/N (S/N = 5) five such spectra were fabricated. For three of those spectra, MKCLASS was unable to determine a spectral type. For the remaining two, MKCLASS classified one as F8 V and the other as G8 V. However, when those S/N = 5 spectra were processed through a low-pass filter (10 point window), MKCLASS succeeded in classifying all of them, returning spectral types between F8 V and G8 V with three in the range G1–G2. For all the other spectra, including the S/N = 10 spectrum, MKCLASS returned accurate spectral types. This performance is comparable to the ability of humans to classify noisy spectra, and may even be slightly superior for S/N $\leqslant$ 20.

4.4. Limitations of MKCLASS

MKCLASS currently has a number of limitations which we list below:

1. Support for classifying O-type stars is currently rudimentary. This capability will be developed once a sufficiently large number of O-type spectra become available for testing.
2. Support for non-MK types, such as carbon stars, white dwarfs, Wolf–Rayet stars, etc., is similarly rudimentary.
3. MKCLASS currently is programmed with spectral criteria only in the 3800–5600 Å region.

5. AVAILABILITY OF MKCLASS

The current version of MKCLASS along with auxiliary programs, two spectral libraries, and documentation may be obtained at the following URL: http://www.appstate.edu/~grayro/mkclass.

REFERENCES

Andrillat, Y., Jaschek, C., & Jaschek, M. 1995, A&AS, 112, 475
Bailer-Jones, C. A. L. 1997, PASP, 109, 932
Carricajo, I., Outeiro, M. M., Rodriguez, A., Dafonte, C., & Arcay, B. 2004, LNEA, 1, 153
Corbally, C. J. 1987, AJ, 94, 161
Gray, R. O. 1989, AJ, 98, 1049
Gray, R. O., & Corbally, C. J. 2009, Stellar Spectral Classification (Princeton, NJ: Princeton Univ. Press)
Gray, R. O., Corbally, C. J., Garrison, R. F., McFadden, M. T., & Robinson, P. E. 2003, AJ, 126, 2048
Gray, R. O., Corbally, C. J., Garrison, R. F., et al. 2006, AJ, 132, 161
Gray, R. O., Napier, M. G., & Winkler, L. I. 2001, AJ, 121, 2148
Gulati, R., Gupta, R., Gothoskar, P., & Khobragade, S. 1994, ApJ, 426, 340
Hanson, M. M., Conti, P. S., & Rieke, M. J. 1996, ApJS, 107, 281
Jackson, P. 1998, Introduction to Expert Systems (Harlow, ND: Pearson)
Kurtz, M. J. 1984, in The MK Process and Stellar Classification, ed. R. F. Garrison (Toronto: David Dunlap Observatory), 136
Kurucz, R. L., Furenlid, I., Brault, J., & Testerman, L. 1984, Solar Flux Atlas from 296 to 1300 nm (Sunspot, NM: Natl. Sol. Obs.)
LaSala, J. 1994, in ASP Conf. Ser. 60, The MK Process at 50 Years, ed. C. J. Corbally, R. O. Gray, & R. F. Garrison (San Francisco, CA: ASP), 312
Manteiga, M., Carricajo, I., Rodríguez, A., DaFonte, C., & Arcay, B. 2009, AJ, 137, 3245
Meyer, M. R., Edwards, S., Hinkle, K. H., & Strom, S. E. 1998, ApJ, 508, 397
Morgan, W. W. 1984, in The MK Process and Stellar Classification, ed. R. F. Garrison (Toronto: David Dunlap Observatory), 18
Morgan, W. W., Keenan, P. C., & Kellman, E. 1943, An Atlas of Stellar Spectra (Chicago, IL: Univ. of Chicago Press)
Singh, H. P., Gulati, R. K., & Gupta, R. 1998, MNRAS, 295, 312
von Hippel, T., Storrie-Lombardi, L. J., & Storrie-Lombardi, M. C. 1994, in ASP Conf. Ser. 60, The MK Process at 50 Years, ed. C. J. Corbally, R. O. Gray, & R. F. Garrison (San Francisco, CA: ASP), 289
Walborn, N. R., Nichols-Bohlin, J., & Panek, R. J. 1985, International Ultraviolet Explorer Atlas of O-Type Spectra from 1200 to 1900 Å (NASA RP 1155; Washington, DC: NASA)
Walborn, N. R., Parker, J. Wm., & Nichols, J. S. 1995, International Ultraviolet Explorer Atlas of B-Type Spectra from 1200 to 1900 Å (NASA RP 1363; Washington, DC: NASA)
Wallace, L., Meyer, M. R., Hinkle, K., & Edwards, S. 2000, ApJ, 535, 325
Wang, S.-G., Su, D.-Q., Chu, Y.-Q., Cui, X., & Wang, Y.-N. 1996, ApOpt, 35, 5155
Weaver, Wm. B. 1994, in ASP Conf. Ser. 60, The MK Process at 50 Years, ed. C. J. Corbally, R. O. Gray, & R. F. Garrison (San Francisco, CA: ASP), 303