

Framework for Tax Preparation

Charles Duan

March 12, 2019

This is a software system for automating the computation and filling of tax forms. Much like commercial software for tax preparation, the object of the system is to allow a user to enter data from received tax forms and other sources, and then answer interview questions, to produce completed tax forms ready for filing.

Unlike commercial tax software, however, this system is not intended to comprehensively cover all tax forms and options. Instead, a second object of this system is to provide a framework to make it reasonably easy for programming-adept users to add tax forms as necessary to their personal situations and needs.

1 Overview of Operation

Use of the system proceeds in three phases.

The first phase is the computation of form values. The user enters relevant data from received tax forms and other sources into files, and then writes a short script that imports that data and computes the values for tax forms to be prepared. This produces a plain text file containing the line item data to be filled into the tax forms.

The second phase is the identification of form fields. For each tax form to be filled in, the user provides the PDF file for the blank form and also the file of line item data. A program displays a graphical interface where the user can click on the blank spaces where each line item should be entered. The result of this is a second plain text data file that lists the coordinates of each of the form fields.

The third phase is the actual filling of the forms. The user writes another short script that receives the data files from the first and second phases,

and places the line items from the first-phase file into the PDF form. If any changes need to be made subsequently, the user may simply rerun the first-phase computation script and then rerun the third-phase filling script, thereby updating all of the tax forms automatically.

2 Phase One: Computation

This section describes the operation of the first phase, computation of tax form line items.

2.1 Structure of Tax Form Data

The basic data structure used throughout the system is the `TaxForm`, which in essence is simply a named table mapping line numbers to values.

3 Phase Two: Form Field Marking

The program that operates this phase is “`mark_fields.rb`”. It must be supplied with the tax form data produced in phase 1, using the “-i” argument. With no further arguments, the program lists the forms for which the fields have not been marked yet.

To mark fields in a form, enter the form name as the first non-option argument to the command. If the form has not been marked yet, you must provide the name of the PDF file for the form as a second command line argument.

The program will automatically save the location of the PDF file and the coordinates of the marked fields to a file, which by default is named “`pos-data.txt`”. This file name may be changed using the “-p” argument.

4 Phase Three: Form Filling

To fill in the forms, create a script that creates a `FormManager` object with the tax forms from phase 1 imported. Then create a `MultiFormManager` to handle the filling of forms. For each form to be filled in, use the `fill_form` method of the `MultiFormManager`.

The MultiFormManager can automatically produce continuation sheets when there is insufficient space. This requires providing two options:

continuation_bio Biographical text for the continuation sheets, just for identification purposes.

continuation_display Either :show, which will print the continuation page to the screen; :raw, which will print the raw troff code; or :append, which will compile the troff code and append the continuation sheet to the filled tax form.

A Forms

A.1 Asset

A form for a capitalized asset.

date The date the asset was put into service.

amount The dollar value of the asset when purchased.

179? Whether the asset is a section 179 deductible asset.

listed? Whether the asset is a listed asset, see Pub. 946 chapter 5.

dc_category The depreciation category for the asset in DC, see the FP-31 instructions, under Depreciation Guidelines.

dc_type Choose from “reference”, “fixed”, or “other” to categorize the asset for DC Form FP-31, lines 1–3.

description A description of the asset.

A.2 Business Expense

A deductible business expense. See IRS Pub. 535, chapter 11.

date The date of the business expense.

amount The dollar amount of the expense.

category The category of expense. “Meals” and “Utilities” will automatically be halved, the former per the IRS 50% rule and the latter on the assumption that 50% of the expense was for non-business purposes. Other common values are “Supplies”, “Travel”, and “Membership”.

description A description of the expense.

A.3 Partner

A single partner within a partnership.

name The partner’s name.

liability Either “general” or “limited”.

nationality Either “domestic” or “foreign”.

type The entity type of the partner. Currently only “Individual” is supported fully. Other possibilities include “Estate”. (This should also, in the future, include corporations, partnerships, trusts, nonprofits, and foreign governments, among others.)

share This partner’s share of profits and losses, as a decimal fraction such that all the partners’ shares add to 1. Currently it is assumed that shares of profits and shares of losses are equal.

capital This partner’s percentage contribution of capital to the partnership, as a decimal fraction such that all the partners’ shares add to 1.

ssn The SSN of the partner.

country The country of citizenship of the partner, relevant to Form 1065 Schedule B-1, part II, column iii.

address The first line of the partner’s address.

address2 The second line of the partner’s address.

A.4 Partnership

Biographical information for a partnership.

name The name of the partnership.

address The first line of the address.

address2 The second line of the address.

business The line of business of the partnership, Form 1065 line A.

product The product or service of the partnership, Form 1065 line B.

code The business code, Form 1065 line C.

ein The employer identification number of the partnership.

start The start date of the partnership.

accounting “Cash”, “Accrual”, or other accounting method.