

Name: \_\_\_\_\_

- This exam is open book, open notes, but there can be no sharing of any material. You can use the Internet, but only as a library. If you are not sure if something is allowed, please ask.
- **COS 480:** Students should answer questions that are not marked with a ★ for 40 points in 40 minutes. Questions marked with a ★ may also be answered *within the 40 minutes*, for extra credit.  
**COS 580:** Students should answer all questions, (★ and non-★) for 60 points in 60 minutes.
- Some questions will use the database instance depicted in Figure 1, which records some information on ferns.<sup>1</sup> (This instance, and the rest of this paragraph, is identical to the description in the homework assignment.) The three title rows of each table indicate attribute names, their data types, and their abbreviated names (for use in relational algebra expressions). The abbreviated name for each table is listed next to its name. The **Ferns** table records the common name, genus, and species of each fern, along with the lower and upper bounds of its typical height range, in feet. The **Habitats** table records the typical habitats for the listed ferns. The **FruitDates** table records the expected fruiting dates (month and day) of the listed ferns.
- Reminder: In all query-writing questions, your answers should work on *all* database instances, not only the instance of Figure 1.
- Show intermediate steps in your answers if you would like your answers to be considered for partial credit in case of errors.

1. (1 pt.) Write your name in the space provided above.

---

<sup>1</sup>The data are derived from (1) George Henry Tilton. *The Fern Lover's Companion: A Guide for the Northeastern States and Canada*. Project Gutenberg, <http://www.gutenberg.org/ebooks/11365>, February 2004 and (2) Anne C. Hallowell and Barbara G. Hallowell. *Fern Finder*. Nature Study Guild Publishers, Rochester, New York, 2nd edition, 2001.

Ferns F

name	genus	species	htLo	htUp
varchar(50)	varchar(25)	varchar(25)	float	float
N	G	S	L	U
Ostrich Fern	Matteuccia	struthiopteris	2	5
Cinnamon Fern	Osmunda	cinnamomea	2	5
Marsh Fern	Thelypteris	palustris	1	2.5
Interrupted Fern	Osmunda	claytoniana	2	4
Rattlesnake Fern	Botrychium	matricariifolium	0.8	2.3

Habitats H

name	habitat
varchar(50)	varchar(25)
N	H
Cinnamon Fern	woods
Cinnamon Fern	wet areas
Ostrich Fern	woods
Ostrich Fern	moist banks
Marsh Fern	fields
Marsh Fern	wet areas
Interrupted Fern	woods
Interrupted Fern	fields
Rattlesnake Fern	woods

FruitDates D

month	day	name
M	D	N
varchar(10)	integer	varchar(50)
May	30	Cinnamon Fern
June	5	Ostrich Fern
June	20	Interrupted Fern
June	25	Rattlesnake Fern
August	30	Marsh Fern

Figure 1: A database of ferns. These tables are repeated on the last page of the test. (You may detach that page and use it for reference. There is no need to reattach it.)

2. (9 pts.) Indicate the result of evaluating each of the following relational algebra (set algebra) expressions on the database instance of Figure 1.

(a)  $\pi_{NHMLU} \sigma_{U < 5} (F \bowtie H \bowtie \sigma_{M = \text{'August'}} D)$

(b)  $\pi_{NM} (H \bowtie D) \bowtie \sigma_{D < 25} \pi_{DLU} (H \bowtie D \bowtie F)$

3. (10 pts.) Indicate the result of evaluating each of the following **extended bag algebra** expressions on the database instance of Figure 1.

(a)  $(\pi_N F - (\pi_N H - \pi_N D)) \cup ((\pi_N F - \pi_N H) - \pi_N)$

(b)  $\tau_{C,N} \gamma_{N, \text{count}(N) \rightarrow C} (\pi_N H \times \pi_D \sigma_{M \neq \text{'June'}} D)$

4. (10 pts.) Write standard SQL queries as directed below.

(a) Write a SQL query to find the heights of the tallest ferns by habitat, using the `htUp` attribute to determine the height of a fern. In more detail, the desired output is a list of tuples  $(h, t)$ , where  $h$  is a habitat and  $t$  is the height of the tallest fern (or ferns) that grows in that habitat. The list should be sorted in descending order of  $t$  and, as a secondary sort, ascending order of  $h$ .

(b) Write a SQL query to find the months in which ferns of at least three different genera<sup>2</sup> fruit. In more detail, the desired output is a set of tuples  $(m)$ , where  $m$  is a month such that there are at least three genera such that there is at least one fern of that genus that fruits in  $m$ .

---

<sup>2</sup>plural of genus.

5. (10 pts.) For each part below, write an extended bag algebra query that is equivalent to the SQL query of the corresponding part of Question 4.

(a)

(b)

6. (10 pts.) ★

- (a) Refer to Question 4a. Instead of only the heights of the tallest ferns by habitat, we now require *both names and heights* of the tallest ferns in each habitat. Write a SQL query for this information, using, as before, the `htUp` attribute to determine the height of a fern. In more detail, the desired output is a list of tuples  $(h, n, t)$ , where  $h$  is a habitat,  $n$  is the name of a fern that grows in  $h$ ,  $t$  is the height of  $f$ , and  $f$  is a tallest fern that grows in  $h$ . Note that there may be more than one tallest fern per habitat; there should be one tuple in the output for each (habitat, tallest-fern) pair. The list should be sorted in descending order of  $t$ , ascending order of  $n$  (secondary sort), and ascending order of  $h$  (ternary sort)..

- (b) Write an extended bag algebra query that is equivalent to the SQL query of Question 6a.

7. (10 pts.) \* Write a relational algebra query that is equivalent to the SQL query of Question 4b **using only the six basic relational algebra operators**: no extended algebra operators, no derived operators, no syntactic sugar (no credit otherwise).

# Scratch page

Material here will not be graded. You may detach and discard this page.

Ferns      F

name varchar(50) N	genus varchar(25) G	species varchar(25) S	htLo float L	htUp float U
Ostrich Fern	Matteuccia	struthiopteris	2	5
Cinnamon Fern	Osmunda	cinnamomea	2	5
Marsh Fern	Thelypteris	palustris	1	2.5
Interrupted Fern	Osmunda	claytoniana	2	4
Rattlesnake Fern	Botrychium	matricariifolium	0.8	2.3

Habitats      H

name varchar(50) N	habitat varchar(25) H
Cinnamon Fern	woods
Cinnamon Fern	wet areas
Ostrich Fern	woods
Ostrich Fern	moist banks
Marsh Fern	fields
Marsh Fern	wet areas
Interrupted Fern	woods
Interrupted Fern	fields
Rattlesnake Fern	woods

FruitDates      D

month M varchar(10)	day D integer	name N varchar(50)
May	30	Cinnamon Fern
June	5	Ostrich Fern
June	20	Interrupted Fern
June	25	Rattlesnake Fern
August	30	Marsh Fern