A Vision: Personal Software Secretaries

Typical tasks

- Manage calendar
- Find useful information on the net
- Conduct routine email business

The Roadblock:

- Requires user-specific customization
- Users will not supply this

Learning apprentice approach:

- Distribute copies of user-independent system
- Each collects training examples of its user's decisions
- Learn user-specific decision making rules
- System evolution: editor $\rightarrow$ assistant $\rightarrow$ representative

Copyright Tom Mitchell, 4 October 1994

Formulating the Learning Problem

Task: Help schedule meetings

Performance: Percent of advice accepted by user

Experience: Observed user scheduling decisions

- How recent? How many? Biased by advice?

Target functions:

Representation of target function:

Learning method:

Copyright Tom Mitchell, 4 October 1994

Thirty-Minute-Meeting?

Attributes:

- DAY-OF-WEEK ACTION CMU-ATTENDEES? DAY-OF-WEEK-OF-LAST-MEETING-WITH-THESE-ATTEN DURATION-OF-NEXT-MEETING-WITH-THESE-ATTENDEES DURATION-OF-LAST-MEETING-WITH-THESE-ATTENDEES GROUP-ATTENDEES? KNOWN-ATTENDEES? MOST-COMMON-DAY-THESE-ATTENDEES MOST-COMMON-DAY-THESE-ATTENDEES-LAST-60-DAYS POSITION-ATTENDEES REQ-COURSE-NAME REQ-EVENT-TYPE REQ-LOCATION REQ-SEMINAR-TYPE SINGLE-PERSON? SPONSOR-ATTENDEES

NO PRUNING:

- on training meetings, 100% coverage, 99% accuracy
- on test meetings: 58 correct, 4 incorrect, 5 not predicted (93% coverage, 94% accuracy)

PRUNING ON TRAINING SET:

- on test meetings: 60 correct, 3 incorrect, 4 not predicted: (94% cov, 95% accuracy)

Copyright Tom Mitchell, 4 October 1994

Learned Tree (Unpruned) On Training Data

```
Learned tree for THIRTY-MINUTE-MEETING?:
(predictions: [correct, incorrect, unknown])
DURATION-OF-NEXT-MEETING-WITH-THESE-ATTENDEES = 60  ==> NO, [8+,0-,0?]
DURATION-OF-NEXT-MEETING-WITH-THESE-ATTENDEES = 90  ==> NO, [21+,0-,0?]
DURATION-OF-NEXT-MEETING-WITH-THESE-ATTENDEES = 30  ==> YES, [44+,0-,0?]
DURATION-OF-NEXT-MEETING-WITH-THESE-ATTENDEES = UNKNOWN,
  REQ-LOCATION = EDRC-CONF-RM  ==> NO, [1+,0-,0?]
  REQ-LOCATION = THANKSGIVING  ==> NO, [2+,0-,0?]
  REQ-LOCATION = WEH1302  ==> NO, [4+,0-,0?]
  REQ-LOCATION = WEH4220,
    DEPARTMENT-ATTENDEES = SCS  ==> YES, [1+,0-,0?]
    DEPARTMENT-ATTENDEES = RI  ==> NO, [1+,0-,0?]
    DEPARTMENT-ATTENDEES = UNKNOWN  ==> NO, [2+,0-,0?]
  REQ-LOCATION = GRAND-CONCOURSE  ==> YES, [1+,0-,0?]
  REQ-LOCATION = WEH7220,
    DAY-OF-WEEK = MONDAY  ==> NO, [1+,0-,0?]
    DAY-OF-WEEK = THURSDAY  ==> YES, [1+,0-,0?]
  REQ-LOCATION = WEH5220  ==> NO, [1+,0-,0?]
  REQ-LOCATION = WEH4601  ==> NO, [5+,0-,0?]
  REQ-LOCATION = MEGHAN-BD  ==> NO, [1+,0-,0?]
  REQ-LOCATION = DC  ==> NO, [1+,0-,0?]
  REQ-LOCATION = WEH5327  ==> YES, [1+,0-,0?]
  REQ-LOCATION = WEH4623,
    DAY-OF-WEEK = MONDAY  ==> YES, [1+,0-,0?]
    DAY-OF-WEEK = FRIDAY  ==> YES, [1+,0-,0?]
    DAY-OF-WEEK = TUESDAY  ==> NO, [7+,0-,0?]
  REQ-LOCATION = RES-HALL-118  ==> NO, [1+,0-,0?]
  REQ-LOCATION = HBH2224  ==> NO, [1+,0-,0?]
  REQ-LOCATION = WEH5409  ==> NO, [3+,0-,0?]
  REQ-LOCATION = HBH1201  ==> NO, [2+,0-,0?]
  REQ-LOCATION = WEH5309,
    POSITION-ATTENDEES = UNKNOWN,
      DURATION-OF-LAST-MEETING-WITH-THESE-ATTENDEES = 60  ==> YES, [2+,0-,0?]
      DURATION-OF-LAST-MEETING-WITH-THESE-ATTENDEES = UNKNOWN  ==> YES, [2+,0-,0?]
      DURATION-OF-LAST-MEETING-WITH-THESE-ATTENDEES = 120  ==> NO, [1+,0-,0?]
    POSITION-ATTENDEES = UNDERGRAD  ==> YES, [2+,0-,0?]
    POSITION-ATTENDEES = FACULTY  ==> NO, [3+,0-,0?]
    POSITION-ATTENDEES = GRAD-STUDENT,
      SINGLE-PERSON? = YES  ==> YES, [4+,0-,0?]
      SINGLE-PERSON? = NO  ==> NO, [1+,0-,0?]
    POSITION-ATTENDEES = ADMINISTRATOR,
      DAY-OF-WEEK = FRIDAY  ==> NO, [1+,0-,0?]
      DAY-OF-WEEK = MONDAY  ==> YES, [1+,0-,0?]
  REQ-LOCATION = WEH4625,
    MOST-COMMON-DAY-THESE-ATTENDEES = MONDAY  ==> NO, [2+,0-,0?]
    MOST-COMMON-DAY-THESE-ATTENDEES = UNKNOWN  ==> YES, [1+,1-,0?]
```

Copyright Tom Mitchell, 4 October 1994

Learned Tree (Unpruned) On Test Data

```
(predictions: [correct, incorrect, unknown])
DURATION-OF-NEXT-MEETING-WITH-THESE-ATTENDEES = 60  ==> NO, [3+,0-,0?]
DURATION-OF-NEXT-MEETING-WITH-THESE-ATTENDEES = 90  ==> NO, [8+,0-,0?]
DURATION-OF-NEXT-MEETING-WITH-THESE-ATTENDEES = 30  ==> YES, [28+,0-,0?]
DURATION-OF-NEXT-MEETING-WITH-THESE-ATTENDEES = UNKNOWN,
  REQ-LOCATION = EDRC-CONF-RM  ==> NO, [1+,0-,0?]
  REQ-LOCATION = THANKSGIVING  ==> NO, [0+,0-,0?]
  REQ-LOCATION = WEH1302  ==> NO, [2+,0-,0?]
  REQ-LOCATION = WEH4220,
    DEPARTMENT-ATTENDEES = SCS  ==> YES, [0+,0-,0?]
    DEPARTMENT-ATTENDEES = RI  ==> NO, [0+,0-,0?]
    DEPARTMENT-ATTENDEES = UNKNOWN  ==> NO, [1+,0-,0?]
  REQ-LOCATION = GRAND-CONCOURSE  ==> YES, [0+,0-,0?]
  REQ-LOCATION = WEH7220,
    DAY-OF-WEEK = MONDAY  ==> NO, [0+,0-,0?]
    DAY-OF-WEEK = THURSDAY  ==> YES, [0+,0-,0?]
  REQ-LOCATION = WEH5220  ==> NO, [0+,0-,0?]
  REQ-LOCATION = WEH4601  ==> NO, [4+,0-,0?]
  REQ-LOCATION = MEGHAN-BD  ==> NO, [0+,0-,0?]
  REQ-LOCATION = DC  ==> NO, [0+,0-,0?]
  REQ-LOCATION = WEH5327  ==> YES, [0+,0-,0?]
  REQ-LOCATION = WEH4623,
    DAY-OF-WEEK = MONDAY  ==> YES, [0+,1-,0?]
    DAY-OF-WEEK = FRIDAY  ==> YES, [0+,0-,0?]
    DAY-OF-WEEK = TUESDAY  ==> NO, [2+,0-,0?]
  REQ-LOCATION = RES-HALL-118  ==> NO, [0+,0-,0?]
  REQ-LOCATION = HBH2224  ==> NO, [0+,0-,0?]
  REQ-LOCATION = WEH5409  ==> NO, [1+,0-,0?]
  REQ-LOCATION = HBH1201  ==> NO, [0+,0-,0?]
  REQ-LOCATION = WEH5309,
    POSITION-ATTENDEES = UNKNOWN,
      DURATION-OF-LAST-MEETING-WITH-THESE-ATTENDEES = 60  ==> YES, [0+,0-,0?]
      DURATION-OF-LAST-MEETING-WITH-THESE-ATTENDEES = UNKNOWN  ==> YES, [0+,1-,0?]
      DURATION-OF-LAST-MEETING-WITH-THESE-ATTENDEES = 120  ==> NO, [0+,0-,0?]
    POSITION-ATTENDEES = UNDERGRAD  ==> YES, [1+,0-,0?]
    POSITION-ATTENDEES = FACULTY  ==> NO, [2+,0-,0?]
    POSITION-ATTENDEES = GRAD-STUDENT,
      SINGLE-PERSON? = YES  ==> YES, [4+,0-,0?]
      SINGLE-PERSON? = NO  ==> NO, [1+,1-,0?]
    POSITION-ATTENDEES = ADMINISTRATOR,
      DAY-OF-WEEK = FRIDAY  ==> NO, [0+,0-,0?]
      DAY-OF-WEEK = MONDAY  ==> YES, [0+,0-,0?]
  REQ-LOCATION = WEH4625,
    MOST-COMMON-DAY-THESE-ATTENDEES = MONDAY  ==> NO, [0+,0-,0?]
    MOST-COMMON-DAY-THESE-ATTENDEES = UNKNOWN  ==> YES, [0+,1-,0?]
```

Copyright Tom Mitchell, 4 October 1994

Some (Postpruned) Rules from the Learned Tree

```
If   The DURATION-OF-NEXT-MEETING-WITH-THESE-ATTENDEES of ?X is 30;
Then The THIRTY-MINUTE-MEETING? of ?X is YES.

If   The SINGLE-PERSON? of ?X is YES, and
     The POSITION-ATTENDEES of ?X is GRAD-STUDENT;
Then The THIRTY-MINUTE-MEETING? of ?X is YES.

If   The DURATION-OF-NEXT-MEETING-WITH-THESE-ATTENDEES of ?X is 90;
Then The THIRTY-MINUTE-MEETING? of ?X is NO.

If   The DAY-OF-WEEK of ?X is MONDAY, and
     The REQ-LOCATION of ?X is WEH5309;
Then The THIRTY-MINUTE-MEETING? of ?X is YES.

If   The REQ-LOCATION of ?X is EDRC-CONF-RM;
Then The THIRTY-MINUTE-MEETING? of ?X is NO.

If   The DEPARTMENT-ATTENDEES of ?X is UNKNOWN, and
     The REQ-LOCATION of ?X is WEH4220;
Then The THIRTY-MINUTE-MEETING? of ?X is NO.

If   The DURATION-OF-NEXT-MEETING-WITH-THESE-ATTENDEES of ?X is 60;
Then The THIRTY-MINUTE-MEETING? of ?X is NO.

If   The DAY-OF-WEEK of ?X is TUESDAY, and
     The REQ-LOCATION of ?X is WEH4623;
Then The THIRTY-MINUTE-MEETING? of ?X is NO.

If   The REQ-LOCATION of ?X is WEH4601;
Then The THIRTY-MINUTE-MEETING? of ?X is NO.

If   The REQ-LOCATION of ?X is WEH1302;
Then The THIRTY-MINUTE-MEETING? of ?X is NO.

If   The POSITION-ATTENDEES of ?X is FACULTY, and
     The REQ-LOCATION of ?X is WEH5309;
Then The THIRTY-MINUTE-MEETING? of ?X is NO.

If   The SINGLE-PERSON? of ?X is NO, and
     The POSITION-ATTENDEES of ?X is GRAD-STUDENT;
Then The THIRTY-MINUTE-MEETING? of ?X is NO.
```

Copyright Tom Mitchell, 4 October 1994

What Next?

Shorten the learner's time contant

- Learn determinations?
  - Seminar location DETERMINED BY Seminar name
- Learn and use domain knowledge?
  - Busier $\rightarrow$ shorter meetings
  - Dean $\rightarrow$ Busier
- Cooperative learning
  - Pool data from multiple users?
  - Learn clusters of users with similar patterns?

Negotiate meetings automatically via email

- Reliability, user trust, user control
- Learn when to ask user permission?

Increase observability of world

- Interface to online information sources
- Even larger hypothesis space...

Copyright Tom Mitchell, 4 October 1994

What Tasks Next?

- Learning more about calendars: when to bump, send reminders, ...
- Learning to conduct routine email negotiations
- Learning to retrieve interesting newsgroup articles
- Learning to navigate through Mosaic

Copyright Tom Mitchell, 4 October 1994