Intelligent Agents

- Woolridge and Jennings
  - agent theory, agent architectures, agent languages
    - weak and strong agents; deliberative and reactive architectures; agent-oriented programming
  - personal digital assistant: e-mail, travel, news articles, etc.
  - softbot: software robot
  - cooperative problem solving: distributed AI
  - interface agents; information agents: collate information from sources to answer question posed by user

- Maes: MIT Media Lab (1997)
  - software agents
  - motivation: number of users, amount of information growing
    - new way of interacting with computer
    - augment direct-manipulation, user responsible for all
  - personalized assistance; autonomous agent
  - various uses: monitoring, critics, etc.

- agents vs. expert systems
  - agents: naive user; common tasks; ok if not completely accurate; personalized; pro-active; adaptive
  - expert systems: expert being assisted; sophisticated tasks; don't want to mislead; more objective; user-driven; fixed knowledge
- key challenges
  - agent needs to know what to do to best assist user
  - interface: how will user communicate with agent

- general approaches
  - knowledge-based approach: give knowledge about user to agents at run-time; sets of rules
  - end-user programming: program by example; rules to follow
  - machine learning: agents tries to program itself; learn from user and from other agents
- problems with other approaches
  - users don't feel they can trust
  - users want personalization
  - users don't like to have to program

- learning agents
  - continuously watch over shoulder of user
  - agents may learn from their peers
  - applications: e-mail, meeting scheduling, recommending web pages, etc.
  - technique of memory-based learning: as you read, record current situational features, compare new situation to memorized one, suggest most similar previous action, with confidence level
  - do-it and tell-me thresholds set to adjust autonomy

- other features of learning agents
  - user can browse what the agent knows
  - user can instruct agent to forget
  - agent can suggest actions to user
  - programmed for single user
- agents asking peers for advice
  - how to deal with multiple conflicting messages
  - modeling trust in peers
  - advice may be sent pro-actively or elicited when desperate

- applications from Maes' group
  - Homer: selects music
  - Webhound: selects webpages
  - Maxims: sorts and filters e-mail
  - Kasbah: buy and sell items
  - Firefly: entertainment recommender matching preferences

- need for agents
  - too much for users to handle
  - limited attention span of users
  - users may not mind giving up control, to save time
  - offloads users from learning tasks (e.g. car repair)

- misconceptions about agents
  - agents are personified: most aren't
  - agents rely on traditional AI inferencing: many use machine learning
  - replaces direct manipulation interfaces: can be complimentary

- challenges in designing effective agents
  - effective interfaces for users to interact with agent
  - requires trust from user, deferring control
    - allow users to see user model being built
    - include explanations
    - vary the level of autonomy
    - allow user programming