

Corso di Ingegneria della Conoscenza

Acquisizione della conoscenza

Andrea Bonarini

Politecnico di Milano



Dipartimento di Elettronica e Informazione
Artificial Intelligence and Robotics Project

Via Ponzio 34/5 - 20133 Milano

Tel. (02) 2399 3525

E-mail: bonarini@elet.polimi.it

Di cosa si tratta?

Acquisire la conoscenza necessaria

perché il KBS risolva il problema

per cui viene realizzato

Diversi tipi di KA con diversi scopi

Acquisizione iniziale

analisi delle caratteristiche di massima delle conoscenze coinvolte in un'applicazione (tipi, sorgenti, quantità)

Acquisizione di base

analisi e modellizzazione completa della conoscenza

Livelli di astrazione nei processi di analisi e modellizzazione

1. Livello linguistico

Identificazione di unità elementari di conoscenza (fatti, oggetti, concetti, fenomeni, eventi, ecc.).

Costruzione di glossari.

2. Livello concettuale

Analisi ed organizzazione delle unità elementari di conoscenza devinute a livello linguistico: identificazione di relazioni. Struttura: proprietà attributi, relazioni, vincoli, ...

Costruzione di modelli utilizzando linguaggi concettuali: reti semantiche, alberi, tabelle, grafi di transizione, ...

3. Livello epistemologico

Integrazione di quanto già individuato in una struttura unitaria ed organica, atta a rappresentare sia le conoscenze, sia i meccanismi di ragionamento che su di esse operano: identificazione della conoscenza strategica e definizione del modello concettuale

Il modello concettuale

Cosa non è

non è una base di conoscenza su carta/calcolatore
non è una rappresentazione intermedia

Cos'è

è una rappresentazione completa, articolata e permanente della struttura della conoscenza di un dominio applicativo sia da un punto di vista statico che dinamico
è una rappresentazione indipendente dall'implementazione
è il risultato principale dell'attività di analisi e modellizzazione della conoscenza

Com'è costruito il modello concettuale

Utilizzando un formalismo per la rappresentazione concettuale:

- espressivo:
permette di esprimere efficacemente tutti i concetti, le relazioni ed i vincoli propri dell'applicazione
- economico:
rappresentazione sintetica, compatta, leggibile e modificabile
- comunicativo:
facilmente comprensibile da tutti gli interessati
- utile:
supporta le attività di analisi e modellizzazione

A cosa serve il modello concettuale

Serve al progettista come base per progettare le strutture di rappresentazione della conoscenza

Serve all'ingegnere della conoscenza come base per l'interpretazione e la codifica della conoscenza

Classi di metodi di acquisizione della conoscenza (1)

TOP-TOWN

Vantaggi:

- disponibilità a priori di raccolta di modelli generici
- poco spazio all'analisi oggettiva e innovativa
- semplice e veloce

Svantaggi:

- rischioso: il modello concettuale potrebbe rivelarsi inadeguato

Classi di metodi di acquisizione della conoscenza (2)

BOTTOM-UP

Vantaggi:

- possibilità di analisi oggettiva e innovativa

Svantaggi:

- lungo e difficile

Possibili problemi per acquisizione della conoscenza

- Metodologie scarse, poco efficaci, poco supportate
- Esperti poco efficaci e poco collaborativi
- Sopravvalutazione delle fonti di conoscenza più accessibili
- Confusione tra relazioni (causali) e relazioni casuali
- Uso di conoscenze e linee di ragionamento semplici e plausibili
- Resistenza ad esplorare conoscenze rare, atipiche, poco frequenti
- Polarizzazione rispetto alle aspettative
- Generalizzazioni errate da esempi poco significativi
- Uso di campioni troppo limitati

Classi di sorgenti di conoscenza

- Esperti:
 - conoscenza palese
 - conoscenza tacita
- Documenti scritti
- Osservazione naturalistica

Tipi di dati: dati verbali

- Incompleti:
 - per omissioni involontarie
 - per reticenza e scarsa cooperatività
 - per difficoltà ad essere esaustivi
 - a causa di conoscenza condivisa non esplicitata
- Non chiari, inaccurati, ambigui
 - intrinsecamente
 - per incapacità di analisi e introspezione
- Scarsamente strutturati e organizzati
- Parzialmente inconsistenti
- Sommari e superficiali
- Spontanei, non controllati o manipolati
- Ricchi di contenuto

Tipi di dati: documenti scritti

- Ragionevolmente completi
- Chiari, accurati, univoci
- Ben strutturati e organizzati
- Consistenti
- Orientati verso uno scopo
- Spesso dettagliati e profondi
- Raramente spontanei, spesso controllati o manipolati
- Non sempre ricchi di contenuto

Tipi di dati: note descrittive

- Obiettive
- Ricche di informazione

Tipi di conoscenza

- Fattuale: fatti, oggetti, fenomeni, eventi del mondo reale. Nomi e struttura.
- Relazionale: proprietà, relazioni, vincoli tra fatti, oggetti, fenomeni, eventi del mondo reale.
- Fenomenologica: regole per proprietà, relazioni, vincoli tra fatti, oggetti, fenomeni, eventi del mondo reale; condizioni che abilitano l'esecuzione di azioni; premesse che portano a conclusioni, ...
- Strategica: modi con cui i diversi tipi di conoscenza possono essere utilizzati per risolvere problemi.

L'intervista degli esperti

Obiettivi:

- raccogliere in modo ordinato le conoscenze palesi
- esplicitare e raccogliere le conoscenze tacite

Problemi:

- efficacia dell'intervista
- tipi e quantità di interviste
- integrazione con altri strumenti di acquisizione

Il piano di intervista

1. Definizione di un piano di intervista

Chi intervistare, con quali scopi, su quali argomenti, con quali tecniche

2. Esecuzione dell'intervista

Acquisizione dei protocolli e loro gestione

3. Analisi dei protocolli

Estrazione della conoscenza, modellizzazione, verifica

4. Revisione ed estensione del piano di intervista

Ripetizione dei punti 2. 3. e 4. finché il modello concettuale non è completato.

Struttura di un piano di intervista

Fase preliminare

Familiarizzazione con il dominio, contatti con utenti ed esperti, pianificazione aggregazione del gruppo di lavoro.

Fase di orientamento

Obiettivo: definire un quadro del dominio ampio, anche se confuso, incoerente, ...

Ciò che importa: imparare a fare le domande giuste (più che sperare di avere buone risposte)

Fase di approfondimento

Obiettivi: esplorazione dei dettagli, eliminazione ambiguità e incoerenze, organizzazione e strutturazione della conoscenza.

Ciò che importa: ottenere le risposte giuste

Fase di completamento e verifica

Obiettivo: costruzione del modello concettuale completo e corretto

Ciò che importa: utilizzare al meglio gli esperti giusti

Fase estensiva

Costruire la base di conoscenza con la qualità desiderata

Ciò che importa: disporre di un buon modello della conoscenza, essere metodici

Tecniche di intervista (1)

Tecniche dirette: ottenere dall'esperto conoscenze che sa esporre direttamente

- tutorial interview
- focused interview
- structured interview
- forward scenario simulation
- thinking aloud
- commentary
- inquisitive observation
- paper review
- teach back review

Tecniche di intervista (2)

Tecniche indirette: ottenere indirettamente, con tecniche opportune, conoscenze che l'esperto non sa esporre direttamente

→ card sorting

→ multidimensional scaling (MDS)

→ repertory grids (RG)

Tutorial interview

Obiettivi:

- stabilire un quadro di riferimento comune
- identificare temi specifici

Modalità:

conversazione informale, aperta, non focalizzata

Risultati:

primi elementi della conoscenza fattuale e relazionale

Note:

tecnica fondamentale per le prime interviste di orientamento

Focused interview

Obiettivi:

→ analizzare temi specifici

Modalità:

→ Conversazione informale, su una lista di al più 4 temi specifici.

→ Per ogni tema il KE si prepara una lista di domande specifiche a cui si attiene in modo flessibile

Risultati:

→ conoscenza fattuale e relazionale

Note:

→ tecnica fondamentale nelle fasi di orientamento ed estensiva

→ meglio scegliere un solo tema per intervista

→ può essere integrata con materiali provenienti da altre fonti

Structured interview

Obiettivi:

→ analizzare una lista di argomenti specifici appartenenti ad un solo tema

Modalità:

→ Interrogazione circoscritta e flessibile su:

- una lista di al più 10 argomenti

- una sequenza logica di al più 15 domande tra loro concatenate

Risultati:

→ raffinamento della conoscenza fattuale e relazionale e fenomenologica

Note:

→ tecnica fondamentale nelle fasi di orientamento ed estensiva

→ meglio scegliere un solo tema per intervista

→ può essere integrata con materiali provenienti da altre fonti

Forward scenario simulation

Obiettivi:

→ analizzare il modo di operare e ragionare dell'esperto

Modalità:

→ l'esperto sviluppa a parole un compito semplice

Risultati:

→ raffinamento della conoscenza fattuale, relazionale e fenomenologica

→ primi elementi della conoscenza strategica

Note:

→ tecnica importante nelle fasi di orientamento ed estensiva

→ mai usarla due volte di seguito

→ il compito da farsi descrivere può venire dall'osservazione naturalistica

Thinking aloud

Obiettivi:

→ analizzare in dettaglio il modo di operare e ragionare dell'esperto

Modalità:

→ racconto libero ad alta voce su ciò che l'esperto fa durante lo svolgimento di un compito scelto da lui

→ diversi livelli di dettaglio: tutto, solo i passi critici, solo ciò che succede ogni X minuti

Risultati:

→ raffinamento della conoscenza fenomenologica e strategica

Note:

→ tecnica importante nelle fasi di completamento e verifica ed estensiva

→ mai usarla due volte di seguito

Commentary

Obiettivi:

→ analizzare in dettaglio il modo di ragionare dell'esperto

Modalità:

→ commento libero dell'esperto, ad alta voce, su ciò che una terza persona fa durante lo svolgimento di un compito scelto dal KE

Risultati:

→ raffinamento della conoscenza fenomenologica e strategica

Note:

→ tecnica importante nelle fasi di completamento e verifica ed estensiva

Inquisitive observation

Obiettivi:

→ analizzare in dettaglio il modo di ragionare dell'esperto

Modalità:

→ interrogazione dell'esperto mentre svolge un compito scelto dal KE

Risultati:

→ raffinamento della conoscenza fenomenologica e strategica

Note:

→ tecnica importante nelle fasi di completamento e verifica ed estensiva

Paper review

Obiettivi:

→ verifica e affinamento della conoscenza

Modalità:

→ raccolta di commenti e critiche degli utenti e degli esperti sui risultati intermedi presentati su supporto cartaceo

Risultati:

→ affinamento della conoscenza fattuale, relazionale, fenomenologica e strategica

Note:

→ tecnica importante nelle fasi di approfondimento e di completamento e verifica

→ un utente alla volta, interessanti diverse opinioni sullo stesso documento

→ noiosa e psicologicamente negativa (critica del lavoro)

Teachback

Obiettivi:

- verifica e affinamento della conoscenza a diversi livelli

Modalità:

- il KE illustra all'esperto un tema, un argomento, una procedura, ... e l'esperto reagisce con conferme e correzioni

Risultati:

- affinamento della conoscenza fattuale, relazionale, fenomenologica e strategica

Note:

- tecnica importante nelle fasi di approfondimento e di completamento e verifica
- non ripeterla più di due volte di seguito

Review

Obiettivi:

- verifica e affinamento della conoscenza a diversi livelli

Modalità:

- raccogliere critiche, commenti, proposte dell'esperto davanti al prototipo in fase di sviluppo, facendogli vedere esempi di comportamento, conoscenze utilizzate, procedimenti seguiti.

Risultati:

- affinamento e verificadella conoscenza fattuale, relazionale, fenomenologica e strategica

Note:

- tecnica fondamentale nella fase esaustiva
- deve essere alternata ad altre tecniche

Card sorting

Obiettivi:

→ analisi della struttura della conoscenza

Modalità:

→ Individuare oggetti, fatti o regole del dominio, scriverli ognuno su un foglietto e chiedere all'esperto di metterli in relazione, ordinandoli o raggrupandoli

Risultati:

→ organizzazione della conoscenza e criteri di organizzazione

Note:

→ tecnica importante nelle fasi iniziali

Card sorting: tecniche specifiche

- dividere le schede in 2 gruppi, a cui devono essere associati nomi; ripetere più volte, rimescolando i foglietti e aumentando il numero dei gruppi
- dividere le schede in gruppi, poi ogni gruppo in sottogruppi, e così via finché non si giunge a gruppi di una scheda; nominare i gruppi; riaggregare le schede; ripetere finché l'esperto non è soddisfatto
- scegliere due schede considerate più simili tra loro, poi aggiungerne un'altra, o scegliere altre 2 schede simili tra loro
- estrarre a caso tre schede e chiedere all'esperto di identificare le due più simili

Multidimensional scaling

Obiettivi:

→ analisi della struttura della conoscenza

Modalità:

→ chiedere all'esperto di valutare la somiglianza tra tutte le coppie di concetti (omogenei) per i quali non esiste una terminologia che descrive le relazioni;
analisi statistica dei valori che permette di trovare le relazioni

Risultati:

→ organizzazione della conoscenza

Note:

→ tecnica importante nelle fasi iniziali

Repertory grid

Obiettivi:

→ analisi della struttura della conoscenza

Modalità:

1. Dialogo iniziale per identificare gli oggetti del dominio
2. Elicitazione delle dimensioni significative del dominio e dei loro estremi
3. Riempimento della griglia: ogni oggetto è inserito nella griglia
4. Clustering degli oggetti: tecniche statistiche che identificano gerarchie
5. Clustering delle dimensioni: matrice di distanza tra le dimensioni

Risultati:

→ organizzazione della conoscenza

Note:

→ tecnica importante nelle fasi iniziali

Problemi del Knowledge Engineer

- essere sopraffatto dallo stato di conoscenza degli esperti e/o dalle esigenze degli utenti
- cedere alla tentazione di sostituirsi all'utente e all'esperto
- essere impaziente
- sentirsi incapace di ottenere ciò che si cerca
- avere difficoltà nell'ottenere cooperazione
- temere il processo di verifica (audit)
- sottovalutare le difficoltà del compito
- sottostimare il problema della polarizzazione e non adottare tecniche per affrontarlo
- preoccuparsi eccessivamente di tempo e soldi
- rinviare la soluzione dei problemi
- confondere le responsabilità delle diverse figure

Regole fondamentali per l'acquisizione della conoscenza

- Non acquisire troppa conoscenza troppo velocemente senza integrarla e verificarla
- Integrare e verificare ogni quanto di conoscenza subito dopo averlo acquisito
- Scegliere attentamente la dimensione dei quanti di conoscenza
- Valutare costantemente la qualità della base di conoscenza in fase di sviluppo