UNIVERSITÀ DI PISA

Corso di Laurea Specialistica in Ingegneria Informatica

CBR
RISK MANAGEMENT SYSTEM

Simone Brienza
Daniele Giannetti

A.A. 2009/2010

Sommario

1 Introduzione 2
1.1 Aspetti generali 2
1.2 Configurabilità 4
1.3 Apprendimento 6
1.4 Limiti del sistema 7
2 Modulo di Gestione della Conoscenza 8
2.1 Rappresentazione degli oggetti in memoria 8
2.2 File di configurazione 22
3 Base di Dati 27
3.1 Tabelle nel database 27
3.2 Vincoli di integrità 30
4 Modulo di Inferenza 32
4.1 Case Based Reasoning (CBR) 32
Fase 1: Estrazione dei progetti simili 34
Fase 2: Ranking dei progetti estratti 39
Fase 3: Identificazione e adattamento dei rischi 42
Fase 4: Identificazione e adattamento delle azioni di mitigazione e recovery 46
Fase 5: Validazione dei risultati del sistema 59
4.2 Chiusura di progetti e aggiunta permanente di casi 63
5 Modulo di Interfaccia 69
6 Case Study: Traduzione dei dati e configurazione del sistema 77
7 Implementazione del sistema 86

1 Introduzione

1.1 Aspetti generali

Si vuole realizzare un sistema intelligente che sfrutti il procedimento di inferenza basato sui casi, ossia un sistema CBR (Case Based Reasoning). Tale sistema verrà utilizzato come sistema di ausilio alle decisioni nel processo aziendale di gestione e controllo dei rischi.

Lo scopo del sistema è duplice:

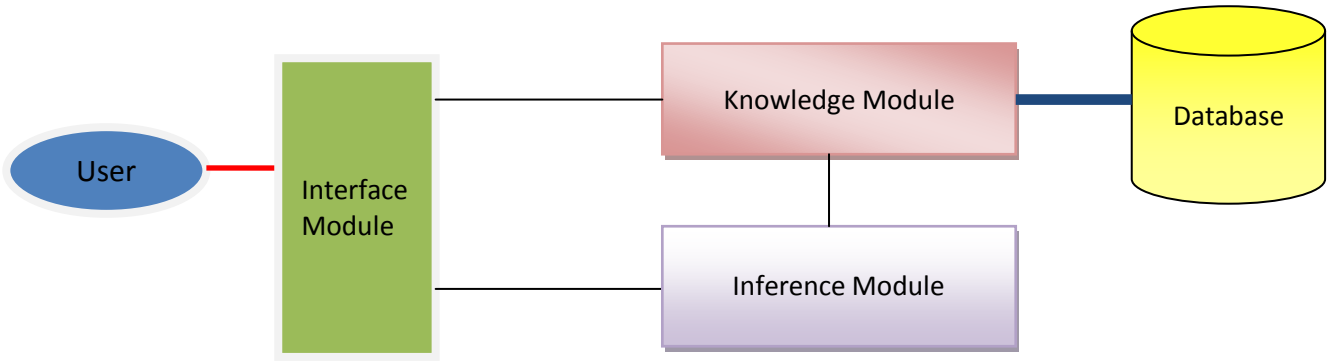
- Dato un nuovo progetto, sfruttando i casi passati presenti in una opportuna base di conoscenza, fornire i rischi che andrebbero gestiti nel progetto assieme alle azioni consigliate, costo potenziale dei rischi, efficacia prevista delle azioni consigliate, costo previsto delle azioni consigliate, ecc…
- Durante lo svolgimento del progetto in azienda, tenere traccia dell'evoluzione dello stesso per assistere il personale nella gestione dei rischi. Inoltre conservare i dati relativi a tutti i progetti (anche se non significativi come casi per le future inferenze) allo scopo di sostituire interamente il precedente sistema di gestione delle informazioni sui progetti.

Nel sistema CBR che si intende realizzare i casi passati su cui si procede con l'inferenza sono i ***progetti passati***, che vengono mantenuti nella base di conoscenza con tutti i parametri che li riguardano, i rischi che si sono affrontati e le azioni ad essi correlate.

Il sistema è composto da tre moduli fondamentali che interagiscono l'uno con l'altro.

a. **Modulo di Gestione della Conoscenza (Knowledge Module)**
   Ha lo scopo di gestire la base di dati e di fornire agli altri moduli funzioni per l'accesso consistente ai dati.
b. **Modulo di Inferenza (Inference Module)**
   Questo modulo, sfruttando i dati raccolti grazie all'interazione col modulo di gestione della conoscenza, elabora le informazioni nella base di dati ragionando sui casi passati (CBR) e genera nuovi risultati adatti al caso in esame.
c. **Modulo di Interfaccia (Interface Module)**
   Il modulo di interfaccia si occupa dell'interazione con l'utente umano, esso cerca di soddisfare le richieste dell'utente sfruttando i servizi offerti dagli altri due moduli. Questo modulo dialoga allora con il modulo di inferenza (qualora al sistema fosse presentato un nuovo progetto da elaborare) e con il modulo di gestione della conoscenza (qualora al sistema venissero richieste delle informazioni su un progetto che già era stato elaborato oppure venissero suggerite dall'esperto umano modifiche da apportare alla base di dati). È essenziale che in questo modulo venga presentata all'utente un'interfaccia grafica intuitiva e user-friendly, in quanto non necessariamente l'esperto umano nel campo di applicazione di questo sistema CBR è anche un esperto di sistemi intelligenti.

La struttura ad alto livello del sistema CBR risulta in definitiva la seguente.



I ruoli dei vari moduli verranno specificati nelle sezioni seguenti, tuttavia notiamo subito alcuni aspetti generali:

- Il modulo di interfaccia dialoga con l'utente tramite una GUI (linea rossa), questo significa che l'interazione avviene mediante pulsanti ed aree di testo. L'interfaccia grafica verrà realizzata tramite il software di sviluppo *Glade*, completamente gratuito.
- L'interazione tra i vari moduli (linee nere) invece avviene tramite l'invocazione di metodi propri del modulo, ad esempio se il modulo di inferenza richiede al modulo di gestione della conoscenza di avere tutti i progetti organizzati in una qualche struttura, questo avviene tramite la chiamata di una funzione.
- L'interazione tra il modulo di gestione della conoscenza e la base di dati stessa (che in pratica è un database *mySQL*) avviene invece tramite l'utilizzo dell'apposita libreria *mysql.h* (linea blu), senza ulteriori complicazioni.

Tutto il software verrà scritto usando il linguaggio Object Oriented *C++.*

Il Database del sistema può essere diviso in due distinte sezioni:

1. *Base di Conoscenza o Base dei Casi*
   Contenente tutte le informazioni riguardanti progetti passati che sono considerati dal sistema come casi passati e dunque utilizzati per l'inferenza sui nuovi progetti forniti dall'utente
2. *Progetti che non costituiscono casi*
   Contenente tutte le informazioni relative a progetti passati che non vengono usati dal sistema nel processo di inferenza, vengono conservati solamente per necessità aziendali e non perché utili in qualche modo al sistema.

Le due sezioni sono mescolate all'interno della base di dati e vengono distinte grazie ad appositi flag (vedere oltre).

1.2 Configurabilità

Il funzionamento del sistema è pienamente *configurabile*, questo significa che l'ambito di applicabilità non è limitato ad un particolare campo di interesse, ma risulta possibile utilizzare dei file di configurazione per la modifica anche profonda del funzionamento del sistema. Questo consente al sistema di adattarsi alla particolare situazione aziendale.

La configurabilità del sistema comporta una serie di conseguenze che verranno affrontate in questa trattazione:

- necessità di organizzare la struttura di un progetto in modo del tutto generico, prevedendo una arbitrarietà dei campi che la compongono;
- Il modulo di inferenza dovrà effettuare la verifica di somiglianza tra progetti operando su parametri del tutto generici in numero, tipo e importanza relativa.

La configurabilità del sistema non si restringe soltanto alla arbitrarietà degli attributi dei progetti su cui lavora, tuttavia questo è sicuramente l'aspetto più significativo. Per effettuare la configurazione degli attributi dei progetti e di altri aspetti significativi, si fa uso di un file apposito scritto dall'utente:

init.conf

In questo file si trovano le seguenti informazioni (vedere oltre per i dettagli sul loro utilizzo):

- Lista degli attributi dei progetti legati al *livello di rischio* del progetto con relativi tipi
- Lista degli attributi dei progetti legati all'*impatto strategico* del progetto con relativi tipi
- Lista dei gruppi in cui sono organizzati gli attributi e appartenenza di ogni parametro ad un singolo gruppo
- Parametri di inferenza configurabili dall'utente, essi vengono utilizzati dal modulo di inferenza e sono i seguenti (vedere oltre per comprendere il loro significato):
  - *K* = costante di contingency
  - *MAX_PROJ_PER_GROUP* = numero massimo di progetti estratti per ogni gruppo
  - *MIN_RI_THRESHOLD* = minimo indice di rischio per un rischio rilevante
  - *MAX_RISKS_PER_PROJ* = massimo numero di rischi che si vuole considerare in ogni progetto
  - *MIN_MID_EFFECTIVENESS* = minima efficacia media percentuale per una buona azione di mitigazione
  - *MAX_MID_INEFFECTIVENESS* = massima inefficacia media percentuale per una azione di mitigazione che non sia troppo dannosa
  - *MOD_THRESHOLD* = soglia di memorizzazione

  Anche se questi parametri vengono passati al sistema inizialmente tramite il file di configurazione init.conf, l'interfaccia grafica permette di modificarli a posteriori per cambiare i criteri di inferenza durante la vita del sistema senza ricostruire la base di dati.

Le informazioni ricavate dal file init.conf vengono memorizzate all'interno della base di dati, ed utilizzate in tutti gli aspetti del funzionamento del sistema.

La struttura (classe) che esprime un progetto ha al suo interno due insiemi di attributi relativi alle due categorie sopra (livello di rischio e impatto strategico), i dettagli della struttura verranno dati nella sezione relativa al modulo di gestione della conoscenza.

L'utilizzo del file di configurazione init.conf da parte del sistema (come anche degli altri file di configurazione) si restringe solo a due casi:

- Primo avvio del sistema
- Comando di rebuild ricevuto dall'utente, che intende dunque ricostruire interamente la base di dati del sistema CBR.

Notare allora che memorizzando tutte le informazioni ricavate dal file init.conf all'interno della base di dati, non è più necessario ai successivi avvii del sistema rifare il parsing di tale documento.
La necessità o meno di effettuare una nuova configurazione iniziale viene dedotta dal sistema in modo semplice esaminando alcune tabelle presenti all'interno della base di dati.
Si assume che non vengano fatte modifiche al database se non tramite le apposite utilità del sistema, ogni modifica fatta al di fuori di quelle previste potrebbe causare il non corretto funzionamento del sistema stesso.
Naturalmente, l'utente può comandare un rebuild della base di dati tramite un controllo da interfaccia grafica, che ha come effetto la cancellazione della base di conoscenza e la riconfigurazione della stessa come in fase di primo avvio.

I parametri configurabili dall'utente utili per l'inferenza sono parametri che possono essere modificati nel corso dell'utilizzo del sistema senza necessità di ricostruire la base di dati, si prevede che alcuni aspetti del funzionamento dell'inferenza (che variano da caso a caso) siano invece richiesti dal modulo di interfaccia direttamente all'utente durante il processo di inferenza (e non vengono dunque specificati in init.conf).

Durante la fase di prima configurazione vengono utilizzati anche altri file, in particolare si prevede di usare:

- file specs.conf
  Per ottenere le specifiche disponibili riguardanti:
  - Quali sono i possibili rischi registrati dal sistema (checklist dei rischi)
  - Quali sono le possibili azioni di mitigazione registrate dal sistema (checklist delle azioni di mitigazione)
  - Quali sono le possibili azioni di recovery registrate dal sistema (checklist delle azioni di recovery)

  Le quali vanno allora a sostituire le checklist aziendali.
- file projects.conf
  Per recuperare i dati sui progetti passati disponibili in azienda e che sono stati registrati prima dell'avvento del sistema CBR di gestione dei rischi.

Ovviamente è necessario fare il parsing dei file nell'ordine in cui sono stati presentati sopra, perché (ad esempio) al momento del recupero delle informazioni sui progetti passati bisogna conoscere quali sono i possibili rischi e le possibili azioni.
Per quanto riguarda le specifiche (primo file), l'utente è in grado una volta completata la prima inizializzazione, di aggiungere rischi e azioni quando vuole tramite l'interfaccia grafica.

La struttura dei tre file di configurazione verrà presentata dopo aver visto la struttura di base con cui i progetti e le specifiche di rischi e azioni vengono rappresentate in memoria, quindi si descriveranno al momento della presentazione del modulo di gestione della conoscenza.

1.3 Apprendimento

Il sistema è in grado di apprendere conoscenza mano a mano che viene utilizzato per l'ausilio nella gestione dei rischi. Un nuovo progetto riceve dal sistema dei rischi e delle azioni ad essi associate sulla base dei casi precedenti. L'esperto umano può poi modificare i risultati in modo che si adattino meglio alla sua esperienza. Al termine del progetto, quando esso viene chiuso, le informazioni possono risultare significativamente diverse da quelle nei progetti già all'interno della base di conoscenza (casi passati).

In particolare, quando un progetto viene presentato al sistema e questo esegue su di esso la valutazione dei rischi e delle azioni si ha che il progetto viene in ogni caso aggiunto alla base di conoscenza come caso significativo. Per questa ragione le inferenze del sistema tengono sempre conto anche dei progetti non ancora chiusi.
Al termine del progetto poi (quando questo viene chiuso) viene valutata l'importanza di mantenere il progetto nella base di conoscenza.
Se il progetto viene ritenuto non significativo, alla chiusura esso viene conservato nel database ma viene attivato un flag che lo marca come non facente parte della base di conoscenza (il che significa che non viene usato per l'inferenza).

Quando si valuta il progetto appena chiuso per capire se vale la pena tenerlo nella base di conoscenza, si fa semplicemente riferimento al numero e alle pesantezza delle modifiche apportate dall'utente al risultato prodotto dal sistema. Se il risultato dell'inferenza viene modificato poco o per nulla dall'utente, la conoscenza in esso contenuta è deducibile dai dati già in possesso del sistema CBR e dunque non costituisce nuova conoscenza da memorizzare. Questo consente anche al sistema di evitare di "tappezzare" lo spazio dei possibili progetti con casi tra loro diversi, ma che comunque non contengono nuova conoscenza utile al processo di inferenza.

Per quanto appena descritto si ha che:

> L'apprendimento del sistema può allora avvenire solamente grazie all'esperto umano che, operando sulla risposta del sistema adattandola alla sua esperienza, produce nuova conoscenza che viene memorizzata dal sistema stesso.

1.4 Limiti del sistema

Concludiamo l'introduzione dicendo che la relazione in esame e l'implementazione proposta sono focalizzati sugli aspetti "intelligenti" del sistema, questo significa che il sistema CBR Risk Management System presenta i seguenti limiti (che possono essere oggetto di futuri sviluppi):

- ✓ Il sistema non consente accesso remoto, si deve prevedere una macchina in azienda dedicata al sistema e accessibile fisicamente dall'esperto umano.
- ✓ Il sistema non supporta accessi concorrenti alla base di dati, come neanche richieste concorrenti pervenenti al modulo di interfaccia, l'utente è allora un unico esperto umano e di ogni modulo del sistema esiste un'unica istanza.
- ✓ Il sistema in esame risulta un software pienamente sequenziale, ossia in pratica abbiamo solo un flusso di esecuzione che si sposta tra i vari moduli in funzione delle chiamate di funzione.
- ✓ Come semplificazione non si prevede la possibilità di comprimere i casi nella base di conoscenza (in modo automatico) tramite creazione di casi artificiali che riassumano l'esito di vari progetti presenti nella base di conoscenza allo scopo di rendere l'inferenza più accurata ed efficiente.
- ✓ I parametri di un particolare progetto non possono venir modificati a posteriori della sua presentazione al sistema, si può solo operare modificando, aggiungendo e rimuovendo azioni di mitigazione/recovery e rischi associati al progetto.
- ✓ Una volta che un progetto viene chiuso dall'utente, esso è memorizzato perennemente nella base di dati e non ulteriormente modificabile (anche se comunque accessibile ed eliminabile tramite interfaccia grafica)

Nelle seguenti sezioni illustreremo la struttura delle varie componenti previste per il funzionamento del software.

2 Modulo di Gestione della Conoscenza

In questa e nelle successive sezioni, pur non entrando nel dettaglio implementativo, si cerca di dare una visione di basso livello sul funzionamento del sistema.

2.1 Rappresentazione degli oggetti in memoria

Ogni oggetto che viene conservato come un insieme di tuple nella base di dati ha una appropriata rappresentazione in memoria, essa viene utilizzata dal sistema per lavorare sull'oggetto stesso. Vediamo solamente i campi principali facenti parte di ogni oggetto, evitando di evidenziare elementi utili solo a scopo implementativo.

Cominciamo dunque con la presentazione della struttura corrispondente ad un progetto (realizzata tramite una classe C++), ogni oggetto di questo tipo dovrà essere efficientemente mappato in tabelle del database relazionale MySQL per la memorizzazione (come vedremo nella sezione relativa alla base di conoscenza).
Tutti i progetti risiedono nella base di dati, quando è necessario recuperare un progetto si procede all'estrazione facendo una copia locale del progetto in memoria. In generale, si cerca di tenere in memoria meno informazioni possibile (entro però dei limiti di complessità di realizzazione).

Come già detto, non tutti i progetti fanno parte della base di conoscenza (vengono utilizzati per l'inferenza come casi passati). Vedremo che nella base di dati del sistema vi è un unico insieme di progetti, ma alcuni di questi possono costituire:

- Progetti copia di altri progetti
- Progetti non contenenti conoscenza utile che non costituiscono dei casi

Un progetto che appartiene ad una delle categorie sopra non viene utilizzato per il processo di inferenza.

Presentiamo allora la struttura di un oggetto Progetto in memoria, dove:

- I campi in nero rappresentano campi sempre presenti e standard del sistema.
- I campi in blu rappresentano campi che l'esperto può configurare tramite il file init.conf in fase di prima accensione, si danno degli esempi.

Project
○ **project_ID** ○ **revision** ○ **state** Assume i seguenti valori possibili: {active, closed} ○ **use_for_inference**[1] ○ **features** • **project_name [S]** • **company_ID [S]** • **customer_name [S]** • **value [N]** • **service_material [S]** • Caratteristiche relative al **livello di rischio** ▪ condizioni del paese in cui il progetto viene eseguito ▪ mercato prospettico per il progetto ▪ comportamento del cliente ▪ forma del contratto col cliente ▪ … • Caratteristiche relative all'**impatto strategico** ▪ importanza del mercato coinvolto per l'azienda ▪ interesse nel mantenere e acquisire il cliente ▪ incidenza sull'immagine aziendale ▪ … ○ Insieme di oggetti di tipo **Proj_Risk** ○ Insieme di oggetti di tipo **Proj_Mit_Action** ○ Insieme di oggetti di tipo **Proj_Rec_Action**

Ogni attributo appartenente all'insieme "Caratteristiche" (Features) deve avere uno dei seguenti cinque tipi (per quelli configurati dall'utente, è possibile decidere il tipo tra i seguenti cinque):

- *[N] Numeric*
  Attributo a valore reale.
- *[E] Enumerate*
  Attributo che assume valore in un insieme discreto di costanti letterali associate a dei naturali e dunque ordinati secondo un qualche criterio direttamente dall'utente.
- *[S] String*
  Attributo che assume come valore una stringa.

[1] Per una descrizione accurata di questo campo si rimanda alla sezione 3 sulla base di conoscenza, in pratica è un campo booleano che descrive se il progetto appartiene o meno alla base di conoscenza (ossia è un caso passato e va usato per l'inferenza).

- *[P] Percentage*
  Attributo che come valore ha un numero reale che deve essere considerato come un valore percentuale ai fini dell'inferenza, nel sistema CBR tutti i valori percentuali sono rappresentati come numeri reali a valori tra 0 e 100 (ad esempio, 5 significa 5% = 0.05).
- *[C] Custom*
  Per attributi di questo tipo l'utente configura sia i *valori possibili*, sia quale è la *prossimità* tra due valori qualunque di questo attributo, tramite una matrice di numeri.
  Per fare un esempio possiamo supporre che l'utente abbia specificato un attributo custom i cui valori possibili sono {Sicilia, Sardegna, Corsica} e abbia specificato la prossimità tra due qualunque valori con una matrice M detta *matrice di somiglianza*, del tipo seguente:

		Attributo nel nuovo progetto		
		Sicilia	Sardegna	Corsica
Attributo nel progetto estratto	Sicilia	100	70	20
	Sardegna	70	100	30
	Corsica	20	30	100

  Notiamo che allora quando si fa il confronto tra due valori di questo attributo, quello del progetto estratto e quello del nuovo progetto, si è sufficiente consultare l'elemento M_{ij} (dove i è il valore nel progetto estratto e j quello nel nuovo progetto).
  Gli elementi all'interno della tabella sono valori percentuali e dunque devono essere specificati dall'utente come numeri da 0 a 100. In condizioni normali la matrice dovrebbe essere simmetrica, ma può anche risultare asimmetrica se l'attributo di tipo custom assume per l'utente un significato particolare.

Per gli attributi predefiniti nell'insieme features (quelli non configurabili dall'utente), sono già stati dati sopra i tipi riportati tra parentesi quadre. Tutti i campi che cadono sotto l'insieme "Caratteristiche"(Features) possono essere usati per il processo di inferenza.

Il company_ID è un attributo predefinito di tipo stringa usato dall'azienda per identificare il progetto, esso non viene però usato dal sistema CBR per tale scopo (per l'identificazione di un progetto è infatti previsto il project_ID, generato dal sistema).

La revisione viene incrementata ogni volta che il progetto subisce una modifica post-validazione (vedere la sezione 5).

Diamo adesso la struttura di un oggetto di tipo Proj_Risk, questi oggetti hanno significato solo nell'ambito di un particolare progetto, per questo possono essere memorizzati direttamente insieme al progetto stesso.

Proj_Risk
○ **risk_ID** (ID delle specifiche del rischio nella base di dati) ○ **company_risk_ID** (campo che identifica il rischio come legato ad un determinato progetto in modo univoco nella base di dati aziendale, si memorizza solo per comodità di utilizzo da parte dell'utente) ○ **strategy** Assume i seguenti valori possibili: {transfer, mitigation, acceptance, mitigation AND acceptance} ○ **state** Assume i seguenti valori possibili (anche combinazioni, vedi oltre): {monitoring, mitigation, recovery, closed_verified, closed_unverified} ○ **revision** (revisione in cui il rischio è stato inserito) ○ **impact** ○ **probability** ○ puntatore al **progetto** contenente l'oggetto di tipo Proj_Risk nella lista dei suoi rischi ○ **description** ○ **causes** (lista delle cause) ○ **effects** (lista dei effetti) ○ Insieme (eventualmente vuoto) di oggetti **Proj_Mit_Bind** ○ puntatore (eventualmente vuoto) ad un <u>unico</u> oggetto **Proj_Rec_Bind** ○ Funzioni per semplificare l'interrogazione dell'utente • Calcolo della contingency (budget allocato per la gestione del rischio) • Calcolo del RI (Risk Index, Indice di Rischio)

I due campi "impact" e "probability" non sono dei semplici numeri, in realtà rappresentano la storia dell'impatto del rischio sul progetto (ossia il costo delle conseguenze qualora il rischio si verifichi) e la storia della probabilità che il rischio si verifichi. Questi due oggetti vengono modificati ad opera della chiusura di azioni di mitigazione o direttamente dall'utente e per questo si tiene traccia dell'evoluzione. Il valore iniziale (all'inizio dell'evoluzione) non può mai essere nullo (a posteriori lo può diventare, ma in tal caso comunque <u>non si può cancellare il rischio</u> in quanto esso rimane necessario per effettuare l'inferenza). Ogni elemento presente in questo storico contiene ovviamente il valore dell'impatto o della probabilità in quella particolare fase della vita del progetto, ma anche un riferimento all'eventuale azione che ha causato la modifica dello storico. La struttura di ogni elemento dello storico (di impatto o probabilità) è la seguente:

- **impact/probability value**
- puntatore ad un oggetto di tipo **Proj_Mit_Action** che rappresenta l'azione che ha portato nello stato descritto dall'elemento stesso nello storico (se è per causa di un'azione che si è aggiunto un elemento allo storico, altrimenti NULL)

Riassumendo allora, nello storico del rischio vengono aggiunti elementi quando:

- un'azione viene eseguita, in tal caso si aggiunge un elemento con il nuovi valore di impatto e un elemento col nuovo valore di probabilità collegati all'oggetto che rappresenta l'azione nell'ambito del progetto;
- l'utente fa una modifica esplicita ad impatto o probabilità, in tal caso si aggiunge un elemento nello storico con il nuovo valore, ma non si lega ad alcuna azione (il puntatore avrà valore NULL).

L'impatto (economico) del rischio può essere sia un valore assoluto (espresso in euro), sia un valore relativo al valore del progetto (dunque espresso come una percentuale), in quale dei due casi ci troviamo dipende unicamente dal valore del campo *rel_risk* interno alle specifiche del rischio (oggetti non relativi al singolo progetto che vengono illustrati più avanti). rel_risk può assumere due diversi valori (è booleano) e in particolare:

- *false* → in questo caso il rischio ha un impatto assoluto indipendente dal valore del progetto (ad esempio il rischio di guasto di un macchinario tipicamente utilizzato, costo di riparazioni e perdita aziendale in termini di tempo macchina noto e fisso).
- *true* → in altri casi un rischio ha un impatto dipendente linearmente dal valore del progetto (ad esempio perdita di un carico di una commessa).

Questo "impatto" nella presente trattazione corrisponde allora a quello che tipicamente in ambito aziendale di risk management si indica con "**costo potenziale impatto**", ossia il costo da sostenere qualora il rischio si verifichi (in euro o percentuale). La gravità economica di un rischio è allora intesa dal sistema direttamente come un costo qualora si verifichi, questa è una differenza con quello che tipicamente si fa in azienda (dove si usa l'"indice impatto" per determinare la gravità di un rischio). È possibile fare una traduzione diretta dell'indice impatto nel costo potenziale impatto, in quanto il significato dei due campi è lo stesso: indicano la pesantezza del costo da sostenere qualora il rischio si verifichi.

Al rischio sono associate delle azioni di mitigazione tramite oggetti del tipo **Proj_Mit_Bind** ed eventualmente un'unica azione di recovery tramite un oggetto **Proj_Rec_Bind**. In particolare, Ogni rischio può venir gestito al più con una unica azione di recovery e questo significa che si sta facendo l'ipotesi che comunque ogni azione di recovery utilizzata per recuperare un certo rischio sia in grado di *recuperare completamente* gli effetti di quel rischio (nel caso si verificassero).

Lo *stato del rischio* si calcola come AND tra le seguenti:

- strategia **mitigazione**:
  - tutte le azioni di mitigazione associate al rischio sono pianificate o chiuse → *monitoring*
  - azioni di mitigazione associate al rischio in corso → *mitigation*
- strategia **accettazione** (azioni di recovery associate al rischio):
  - *recovery*

Vedremo meglio più avanti come avviene il passaggio dello stato del rischio a "closed_verified" e a "closed_unverified", e quale è il loro significato.

Nel caso in cui come strategia sia stato scelto lo spostamento per un rischio (transfer), non si prevedono azioni associate al rischio, anche se il rischio in effetti esiste. Lo stato del rischio in questo caso è sempre monitoring.

Il puntatore, presente in ogni oggetto di tipo Proj_Risk che punta all'oggetto di tipo Project contenente il rischio è utile in varie occasioni, prima tra tutte quando si vuole calcolare la contingency associata al rischio tramite la funzione di utilità (la quale può aver bisogno del valore economico del progetto).

Per quanto riguarda le funzioni di utilità notiamo che:

- Il calcolo della contingency di un rischio avviene come segue:

  $$cont_rischio = \sum cost_mit[euro] + (risk_contingency[euro])_{if\ acceptance\ \ in\ strategy}$$

  Dove in pratica si fa la somma del costo di tutte le azioni di mitigazione associate al rischio e la parte di contingency dell'azione di recovery (se prevista) necessaria per il particolare rischio in esame, vedremo oltre dove si trovano effettivamente la risk_contingency e il costo delle azioni di mitigazione.
- L'indice del rischio (RI) si calcola semplicemente come:

  $$RI = impact[in\ euro] * \frac{probability}{100}$$

  (con impatto e probabilità parametri attuali nello storico)

  Essendo che si opera su parametri attuali, ovviamente il RI cambierà durante la gestione del rischio, esso fondamentalmente rappresenta la <u>priorità del rischio</u>. Notare che tipicamente in azienda non si usa l'impatto in euro (costo potenziale impatto) ma l'indice impatto (un intero che rappresenta la gravità del verificarsi del rischio. Essendo le due grandezze legate e sostanzialmente proporzionali, per il calcolo della priorità del rischio si considera l'impatto in euro.

  Nel caso in cui rel_risk sia false, l'impatto in euro si legge direttamente nello storico, altrimenti si deve ricavare dal valore del progetto usando la percentuale memorizzata nello storico come segue (se rel_risk è true):

  $$impact[in\ euro] = value * \frac{impact}{100}$$

Vediamo adesso la struttura di un oggetto di tipo Proj_Mit_Bind, questi hanno valore nell'ambito di un progetto e sono legati al rischio come visto sopra.

Proj_Mit_Bind
○ puntatore ad un oggetto **Proj_Mit_Action** che individua l'azione di mitigazione ○ **impact_effectiveness** ○ **probability_effectiveness**

L'efficacia in termini di impatto e probabilità dell'azione puntata sul rischio (impact_effectiveness e probability_effectiveness) sono sempre valori <u>percentuali</u> che indicano la *percentuale del valore di probabilità o impatto del rischio che bisogna andare a sottrarre al valore attuale (nello storico) dopo che si è applicata l'azione*. In ogni caso questi campi possono avere valore sotto lo zero in quanto un'azione può anche avere effetti negativi per la mitigazione di un rischio.

Notiamo che il valore di efficacia in termini di impatto resta consistente qualunque sia il tipo del rischio (relativo o assoluto), possiamo fare due esempi:

- Rischio con impatto attuale del 7% (rel_risk = true) a cui viene applicata un'azione che ha efficacia in termini di impatto del 20%. Impatto post-mitigazione del rischio pari a:

$$new_impact\ = 7 - 7 * \frac{20}{100} = 7 - 1.4 = 5.6\%$$

- Rischio con impatto attuale di 2700 (rel_risk = false) a cui viene applicata un'azione che ha efficacia in termini di impatto del 20%. Impatto post-mitigazione del rischio pari a:

$$new_impact\ = 2700 - 2700 * \frac{20}{100} = 2700 - 540 = 2160\ euro$$

Quanto sopra illustra dunque come l'efficacia in termini di impatto va effettivamente ad agire sullo storico dell'impatto del rischio aggiungendo elementi alla sua storia, per quanto riguarda l'efficacia in termini di probabilità la situazione è analoga:

- Rischio con probabilità attuale del 18% a cui viene applicata un'azione che ha efficacia in termini di probabilità del 40%. Probabilità post-mitigazione del rischio pari a:

$$new_probability\ = 18 - 18 * \frac{40}{100} = 18 - 7.2 = 10.8\%$$

Si potrebbe a questo punto pensare che l'ordine in cui vengono applicate le azioni di mitigazione influisca sul risultato finale di impatto e probabilità del rischio nello storico, è semplice dimostrare che non è così:

- Supponiamo di avere un'azione di mitigazione A_1 e un'azione di mitigazione A_2, facciamo i conti con la probabilità anche se essi sono facilmente estendibili all'impatto.
- Sia la probabilità iniziale del rischio P_i, quella finale P_f e P_m la probabilità intermedia dopo l'applicazione di una delle due azioni.
- Sia E_1 l'efficacia dell'azione A_1 in termini di probabilità ed E_2 l'efficacia dell'azione A_2.

Vediamo che succede applicando le azioni nei due ordini possibili.

1. A_1 poi A_2

$$P_m = P_i - P_i \frac{E_1}{100} = P_i * \left(1 - \frac{E_1}{100}\right)$$

$$P_f = P_m - P_m * \frac{E_2}{100} = P_m * \left(1 - \frac{E_2}{100}\right) = P_i * \left(1 - \frac{E_1}{100}\right) * \left(1 - \frac{E_2}{100}\right)$$

2. A_2 poi A_1

$$P_m = P_i - P_i * \frac{E_2}{100} = P_i * \left(1 - \frac{E_2}{100}\right)$$

$$P_f = P_m - P_m * \frac{E_1}{100} = P_m * \left(1 - \frac{E_1}{100}\right) = P_i * \left(1 - \frac{E_2}{100}\right) * \left(1 - \frac{E_1}{100}\right)$$

Questo significa che l'ordine in cui si applicano le azioni di mitigazione non è importante per il risultato finale, e che l'effettivo apporto dell'azione di mitigazione non cambia se essa viene svolta prima o dopo.

Per questo risulta legittimo basare l'inferenza (come vedremo successivamente) sui valori presenti nei campi di efficacia degli elementi Proj_Mit_Bind nella base di conoscenza, senza curarsi della possibilità che l'azione sia stata inserita dall'utente solo a posteriori dell'esecuzione di altre azioni di mitigazione[2].

La struttura si un oggetto di tipo Proj_Rec_Bind è invece la seguente:

Proj_Rec_Bind
○ puntatore ad un oggetto **Proj_Rec_Action** che individua l'azione di recovery ○ **risk_contingency** ○ **original_RI**

Ad un'azione di recovery è associata una contingency per ogni rischio a cui essa è collegata (intesa come soldi da mettere da parte per il recupero del rischio), la contingency per un particolare rischio è contenuta nell'elemento Proj_Rec_Bind che lega il rischio all'azione di recovery (campo risk_contingency).

È preferibile non memorizzare direttamente il valore di questa contingency nell'azione di recovery nell'ambito del progetto, si sceglie invece tenere memoria della contingency relativa ad ogni rischio gestito dall'azione di recovery sul legame tra il rischio e l'azione stessa.

Il denaro da mettere da parte per la gestione di un rischio ad opera di una azione di recovery tipicamente si aggira attorno ad un valore calcolabile a partire dai parametri del rischio nel momento in cui l'azione di recovery viene aggiunta, per cui si definisce *valore di partenza* del campo risk_contingency la quantità:

$$VPart = \frac{probability}{100} * impact * K$$

[2] Questo sempre assumendo comunque che l'efficacia delle azioni di mitigazione non cambi a seconda che nel progetto siano state o meno già eseguite altre azioni di mitigazione, ossia non si considera per semplicità l'eventuale correlazione tra diverse azioni di mitigazione.

Dove K (*costante di contingency*) è uno dei parametri di inferenza configurabili dall'utente compreso tra 0 e 1, probability e impact sono probabilità e impatto del rischio nel momento in cui l'azione è stata aggiunta per la gestione del rischio.
Con *valore di partenza* si intende un valore più o meno tipico per le contingency in azienda, questo viene usato nel processo di inferenza per capire quali azioni di recovery si sono comportate meglio e quali peggio del valore tipico (per privilegiare alcune azioni di recovery su altre).
È ovvio che, data la risk_contingency associata a tutti i rischi dove una certa azione di recovery è stata prevista come recupero, è possibile ricavare la contingency totale per l'azione di recovery come segue:

$$contingency = \sum risk_contingency$$

Interessante è notare che il valore di partenza usato per l'inferenza del campo risk_contingency dipende dall'impatto (al momento di inserimento dell'azione di recovery), il quale assume due diversi significati a seconda del valore di rel_risk. Dunque anche risk_contingency può assumere due diversi significati. Questo significa che (anche se l'utente opera sempre sul valore di risk_contingency lavorando in euro) il sistema si occuperà di convertire internamente il campo risk_contingency dividendo per il valore del progetto qualora rel_risk per il rischio sia a valore true.

Per calcolare il valore di partenza del campo risk_contingency in fase di inferenza è necessario conoscere impatto e probabilità del rischio nel momento in cui l'azione è stata aggiunta. Si memorizza allora direttamente all'interno della struttura che lega il rischio all'azione di recovery (campo original_RI) il RI (indice di rischio) come calcolato nel momento in cui l'azione di recovery è stata aggiunta. $original_RI * K$ ci darà il valore di $VPart$ necessario per l'inferenza.
Avendo definito precedentemente il RI usando l'impatto in euro, ne deriva che original_RI non corrisponde alla definizione precedente in quanto si utilizza l'impatto indipendentemente dal particolare valore di rel_risk (dunque anche original_RI assume due diversi significati in funzione di rel_risk).

$$original_RI = \frac{probability}{100} * impact$$

Il motivo per cui è necessario poter confrontare il valore in risk_contingency con un valore di partenza dato da impatto e probabilità al momento dell'inserimento dell'azione di recovery deriva dal fatto che all'utente è consentito l'inserimento di azioni di recovery anche a posteriori di una prima mitigazione del rischio (come vedremo più avanti). Quando l'utente fa un inserimento di quel tipo si baserà sui valori attuali di probabilità e impatto del rischio, e noi bisogna tenerlo presente per capire se effettivamente un'azione di recovery è stata più o meno efficace della media aziendale (ossia se è costata meno della media).

Anche il campo risk_contingency, come già detto, assume un significato diverso a seconda del valore di rel_risk:

- Se l'impatto del rischio *non* è relativo allora risk_contingency equivale al valore numerico della somma di denaro in euro che è necessario mettere da parte.
- Se l'impatto del rischio è relativo invece risk_contingency assume un valore numerico che è dato dalla somma di denaro in euro da mettere da parte divisa per il valore del progetto e moltiplicata per 100 (ossia si memorizza il valore percentuale). Questo si fa perché così facendo il valore in risk_contingency è direttamente confrontabile con original_RI*K, e questo è utile per il processo di inferenza.

Diamo adesso la struttura di un oggetto di tipo Proj_Mit_Action, questi oggetti hanno un significato solo nell'ambito di un particolare progetto, per questo compaiono in memoria direttamente insieme al progetto stesso.

Proj_Mit_Action
○ **mitigation_ID** (ID delle specifiche dell'azione di mitigazione nel database) ○ **cost** ○ **description** ○ **revision** ○ **state** Assume i seguenti valori possibili: {planned, active, closed}

Il valore del campo costo (*cost*) può essere in euro oppure un valore percentuale riferito al valore del progetto, dipende da un campo *rel_act* definito all'interno dell'oggetto di tipo specifica dell'azione di mitigazione (vedere oltre):

- *true* → l'azione di mitigazione ha un costo relativo al valore del progetto (espresso in percentuale tra 0 e 100), ad esempio se il progetto si occupa di trasporto di vetro allora l'acquisto di nuovo imballaggio ha un costo relativo al valore del progetto.
- *false* → l'azione di mitigazione ha un costo assoluto (espresso in euro), ad esempio se il progetto necessita come mitigazione di determinati rischi di assumere un esperto che si occupi di alcuni aspetti tecnici non noti in azienda (o per l'addestramento del personale) .

Il campo *revision* è inizialmente vuoto e viene lasciato tale fino a che l'azione di mitigazione non passa effettivamente a stato chiuso. Quando l'azione di mitigazione si chiude (è stata completamente eseguita) la revisione del progetto non viene incrementata e in questo campo viene salvato il valore della revisione attuale del progetto (si registra in questo modo che l'azione di mitigazione ha avuto il suo effetto in quella revisione).

Possiamo adesso fornire i dettagli della *chiusura di azioni di mitigazione e di rischi*
La chiusura di un'azione di mitigazione viene lanciata dall'utente esplicitamente in quanto si vuole segnalare al sistema gli effetti di un'azione che è stata eseguita.

Si presenta all'utente una schermata contenente tutti i rischi associati al progetto su cui l'azione doveva avere un effetto da previsioni, con i relativi effetti previsti. L'utente può allora:

- cancellare un legame tra l'azione ed un rischio in quanto non c'è stato alcun effetto.
- modificare i valori per l'efficacia dell'azione su un qualunque rischio prima della chiusura.
- modificare il costo dell'azione di mitigazione (il costo è risultato diverso da quello atteso).

Ogni volta che si opera la chiusura di un'azione di mitigazione si vanno a scrivere i nuovi valori per impatto e probabilità nello storico contenuto nei rischi su cui tale azione ha avuto un effetto.

Lo stato di un rischio invece passa a **"closed_unverified"** quando l'utente intende segnalare che il rischio in questione non è più da considerare nel risk management del particolare progetto (qualsiasi sia il motivo) anche se i suoi effetti *non si sono ancora verificati*. Quando l'utente compie questa procedura il rischio non viene eliminato dal progetto perché si ritiene che la sua presenza sia significativa e utile per l'inferenza su futuri progetti. Tutte le azioni legate al rischio vengono conservate nella base di conoscenza ed anche i legami col suddetto rischio. L'utente ha comunque la libertà di modificare l'efficacia delle azioni sul rischio, ed eventualmente di eliminarle. Vi sono alcune differenze rispetto a lasciare il rischio non in uno stato "closed" che rendono questo stato fondamentale:

- Un rischio closed_unverified non viene mostrato all'utente nella lista dei rischi da gestire per il progetto.
- Azioni di mitigazione in chiusura non provocano alcun aggiornamento allo storico del rischio, anche se legate allo stesso con una certa efficacia. Questo significa che nello storico del rischio verranno trovati gli effetti di azioni chiuse prima del passaggio del rischio in stato "closed_unverified", ma non quelli delle azioni di mitigazione chiuse dopo che il rischio ha cambiato stato. È ovvio che, se un rischio passa in stato "closed_unverified", le azioni che vengono chiuse dopo la chiusura del rischio non dovrebbero avere lo stesso peso di quelle chiuse prima (infatti sono azioni di cui l'utente non conosce l'efficacia effettiva al momento della chiusura), per questo in fase di inferenza risulta utile esaminare lo storico del rischio per capire quando l'azione è stata chiusa.

Lo stato del rischio passa a **"closed_verified"** quando l'utente intende segnalare che la gestione del rischio si è conclusa perché gli effetti del rischio si sono verificati nel progetto.
In questo caso si assume che ogni azione di mitigazione associata al rischio ma non ancora in stato chiuso non sia stata completata e non abbia generato effetti sul rischio, comunque si procede come sopra mantenendo le azioni di mitigazione previste nel progetto per il management del rischio in questione. Se questi legami non vengono eliminati esplicitamente dall'utente, il sistema continuerà a considerare tali legami come conoscenza utile per l'inferenza. Questo rischio verrà sempre mostrato all'utente come rischio verificato, e come sopra azioni di mitigazione in chiusura dopo la chiusura del rischio non provocano alcun aggiornamento nello storico del rischio, anche se legate allo stesso con una certa efficacia. Si opera come sopra per capire quando un'azione è stata chiusa, sempre allo scopo di evitare di dare troppa importanza ad azioni che in realtà non sono mai state eseguite nell'ottica del particolare rischio.

Se viene chiuso il progetto di conseguenza non si modifica lo stato delle azioni di mitigazione all'interno del progetto, queste rimarranno congelate all'interno della base di conoscenza. Il fatto che non si provveda alla chiusura delle azioni di mitigazione è positivo perché le azioni che non hanno ancora subito un feedback sono considerate successivamente meno affidabili (per l'inferenza). Una volta che un progetto è stato chiuso *non è più possibile operare sulle sue informazioni*. Tutti i rischi non closed presenti nel progetto vengono *forzati in stato "closed_unverified"*.

Notare allora che il sistema non opera mai in automatico l'operazione di chiusura di un'azione in quanto non può sapere se un'azione di mitigazione è stata effettivamente conclusa.

A questo punto presentiamo anche la struttura di un oggetto di tipo Proj_Rec_Action, anche queste hanno significato solamente nell'ambito di un particolare oggetto Project.

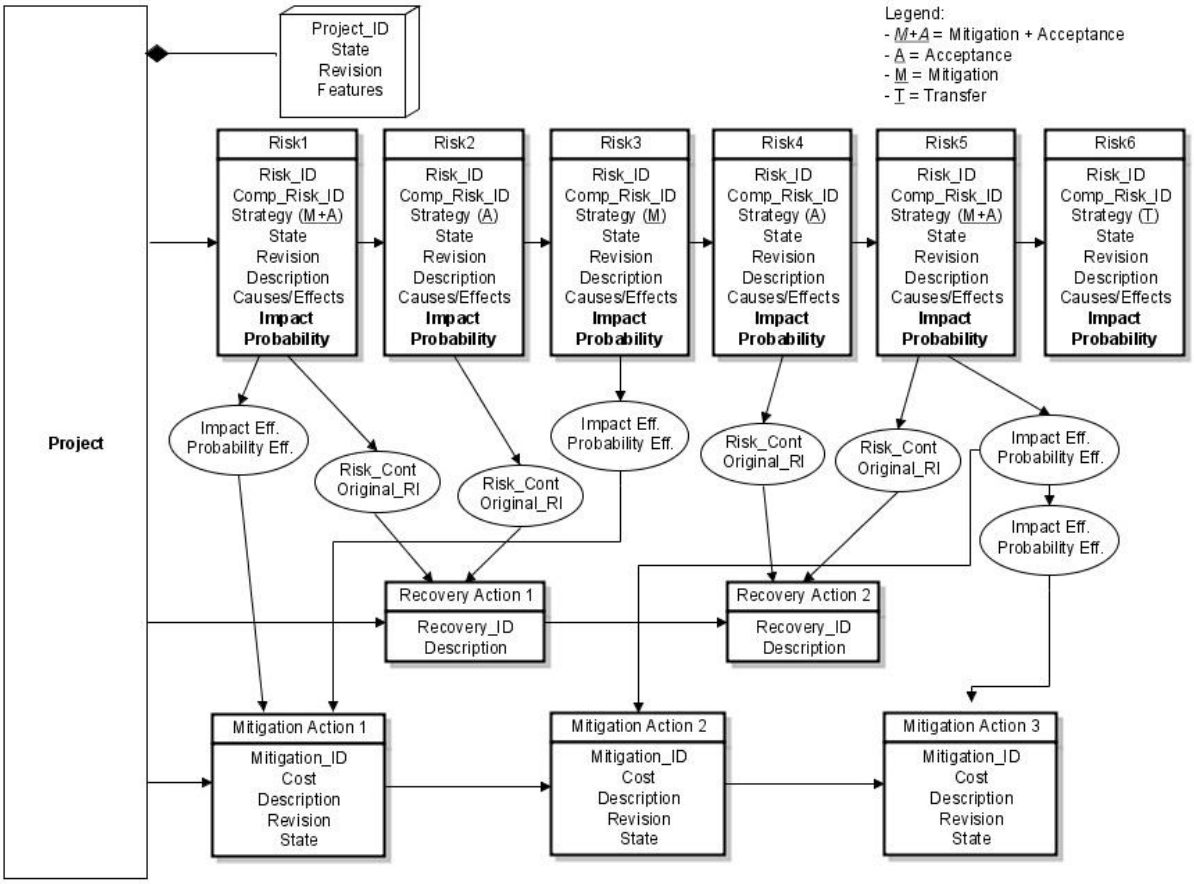
Proj_Rec_Action
○ **recovery_ID** (ID delle specifiche dell'azione di recovery nel database) ○ **description**

Il motivo per cui *non* si è deciso di inserire i valori di original_RI all'interno di questa struttura al posto che sulle connessioni (Proj_Rec_Bind) è duplice:

- Un'azione di recovery può gestire diversi rischi, dunque molteplici original_RI andrebbero mantenuti in una struttura dinamica;
- Anche se l'azione è aggiunta in un dato istante, è sempre possibile che essa sia stata configurata per il recovery di un rischio in un istante successivo e dunque in realtà si vuole tener traccia del RI che si aveva *quando l'azione di recovery è stata legata al rischio*.

Per le azioni di recovery si assume che, se l'azione non viene scollegata esplicitamente da un rischio prima che esso si chiuda, il rischio sia comunque recuperato nel caso si verifichi (le azioni di recovery *non hanno stato*).

In definitiva possiamo allora descrivere la struttura di un oggetto di tipo progetto con un semplice diagramma esemplificativo, non si usa una notazione standard come un diagramma UML, ma l'oggetto risulta molto intuitivo:



A questo punto terminiamo la trattazione dei dati interni al modulo di gestione della conoscenza come visti in memoria con la presentazione delle tre classi specifica del rischio (Risk), specifica di un'azione di mitigazione (Mit_Action) e specifica di un'azione di recovery (Rec_Action). Queste tre sostanzialmente corrispondono ad oggetti nelle checklist aziendali dei rischi e delle azioni.

Notare che queste tre classi descrivono oggetti completamente indipendenti da un progetto inteso come oggetto, questo significa che nella base di dati verranno memorizzate a parte, inoltre non è richiesta l'esistenza di un oggetto progetto a cui queste classi fanno riferimento.

Risk
○ **risk_ID** ○ **risk_CL_code** (campo che identifica il particolare rischio nella checklist aziendale dei rischi, fondamentalmente è previsto solo per comodità dell'utente e non è utilizzato dall'inferenza) ○ **category** ○ **description** ○ **causes** (lista delle possibili cause) ○ **effects** (lista dei possibili effetti) ○ **rel_risk**

Mit_Action
○ **mitigation_ID** ○ **action_CL_code** (campo che identifica la particolare azione nella checklist aziendale delle azioni, fondamentalmente è previsto solo per comodità dell'utente e non è utilizzato dall'inferenza) ○ **rel_act** ○ **description**

Rec_Action
○ **recovery_ID** ○ **action_CL_code** (campo che identifica la particolare azione nella checklist aziendale per le azioni, fondamentalmente è previsto solo per comodità dell'utente e non è utilizzato dall'inferenza) ○ **description**

Notare che è presente una descrizione sia all'interno delle specifiche per i rischi/azioni di mitigazione e recovery appena descritte, sia all'interno delle particolari istanze di rischio e azione all'interno dei progetti gestiti dal sistema (Proj_Risk, Proj_Mit_Action e Proj_Rec_Action). Questo è fatto per concedere all'utente un grado di libertà aggiuntivo: quando il sistema ricava determinati rischi e azioni per un nuovo progetto fornito dall'utente, la descrizione per il rischio o l'azione nel particolare progetto coincide inizialmente con quella presente nelle specifiche appena viste, tuttavia l'utente può modificare a posteriori questa descrizione specializzandola per il particolare caso (progetto) in esame.

Si deve inoltre sottolineare che anche i campi "causes" e "effects" sono ripetuti sia nelle specifiche per il rischio sia nelle particolari istanze di rischio presenti nei vari oggetti progetto. Questo è fatto per motivi simili a quanto detto per il campo description: si da un grado di libertà in più all'utente consentendogli di specializzare la lista delle cause e degli effetti nell'ambito del particolare progetto. In generale si assume che le cause ed effetti prese in considerazione nell'ambito di un particolare progetto saranno un sottoinsieme di quelle presenti nelle specifiche del rischio.

Vediamo la struttura del file di configurazione *init.conf*, tenendo presente il formato con cui viene rappresentato un progetto in memoria. Non abbiamo ancora visto nei dettagli il significato di tutti i parametri utili per l'inferenza che vanno specificati in questo file, essi verranno esaminati in sezione 4. Per questo e tutti i seguenti file di configurazione vi sono delle convenzioni sintattiche:

- '%' corrisponde all'inizio di un commento che dura fino a fine riga, il testo che risulta parte di un commento è stato scritto nei file di esempio in colore blu.
- <sez> indica l'inizio di una sezione di tipo "sez"
- [sottosez] indica l'inizio di una sezione di tipo sottosez contenuta all'interno di una sezione descritta tramite parentesi angolari.

La descrizione dei file di configurazione non viene data in una specifica grammaticale rigorosa, si usa invece una presentazione basata su esempi. Per il corretto funzionamento del sistema è necessario che l'utente non utilizzi all'interno dei file di configurazione come parte dei nomi e di altre stringhe i simboli '%', '[', ']', ':', '|', ',', ';', '<', '>' e ' ' ' (i quali assumono un significato particolare nel parsing dei file di configurazione e vanno usati solo con tale significato). Per quanto riguarda le stringhe che rappresentano descrizione, cause ed effetti di rischi ed azioni è consentito anche l'utilizzo di ' ' '.

```
% file init.conf

% attributi legati al livello di rischio
<RISK_LEVEL>
  % attributo numerico (N) misurato in km di nome lunghezza
  lunghezza : N : km
  citta : S % attributo stringa (S) di nome citta
  % attributo enumerato di nome mercato con possibili valori:
  % molto scarso, scarso, medio, buono, ottimo
  mercato : E : molto scarso | scarso | medio | buono | ottimo
  % l'ordine è importante nel caso sopra, ma può anche essere inverso,
  % ad esempio ottimo, buono, medio, scarso, molto scarso
  % attributo percentuale di nome gradimento
  gradimento : P
  % attributo custom di nome isola con matrice di somiglianza
  isola : C : Sicilia | Sardegna | Corsica [ 100 70 20; 70 100 30; 20 30 100 ]
  % si specifica riga per riga la matrice di  somiglianza
  sicurezza : P
  % altri attributi risk level ...
  % ...

% attributi legati all'impatto strategico del progetto
<STRATEGIC_IMPACT>
  % attributo numerico (N) misurato in euro di nome finanziamento
  finanziamento: N : euro
  % attributo stringa (S) di nome partner
  partner : S
```

```
  % attributo enumerato di nome produttività con possibili valori:
  % bassa, medio bassa, buona, ottima
  produttivita : E : bassa | medio bassa | buona | ottima
  % l'ordine è importante nel caso sopra, ma può anche essere inverso,
  % ad esempio ottima, buona, medio bassa, bassa
  % attributo percentuale di nome probabilità di fallimento
  probabilita di fallimento : P
  % attributo custom di nome imp_cliente con matrice di somiglianza
  imp_cliente : C : Gold | Silver | Bronze | Casual [ 100 50 30 20; 50 100 70 50; 30 70 100 70; 20 50 70 100 ]
  % si specifica riga per riga la matrice di somiglianza
  % altri attributi strategic impact
  % ...

% gruppi definiti per gli attributi
<GROUPS>
  % gruppo di nome tecnico che raggruppa attributi che rappresentano
  % aspetti tecnici del progetto
  tecnico : lunghezza
  % gruppo di nome economico che raggruppa attributi che
  % rappresentano l'ambito economico del progetto
  economico : mercato, gradimento, partner, finanziamento, isola
  % altri gruppi che raggruppano attributi definiti sopra,
  % ogni attributo appartiene ad uno e un solo gruppo
  % ...

% parametri di inferenza che devono essere configurati dall'utente,
% questi potranno venire modificati anche a posteriori della prima
% configurazione (tramite GUI), il loro nome è fissato e
% l'utente deve definire tutti e soli questi attributi in questo file
% per la prima configurazione.
<INFERENCE_PARAMETERS>
  % esempi dei vari parametri di inferenza, l'ordine è arbitrario
  K = 0.7 % costante di contingency
  MAX_PROJ_PER_GROUP = 15 % numero massimo di progetti estratti per ogni gruppo
  MIN_RI_THRESHOLD = 2000 % minimo indice di rischio perché un rischio sia rilevante in estrazione
  MAX_RISKS_PER_PROJ = 50 % massimo numero di rischi che si vuole considerare in ogni progetto
  MIN_MID_EFFECTIVENESS = 10 % minima efficacia media percentuale per una buona azione di mitigazione (es 10%)
  MAX_MID_INEFFECTIVENESS = 2 % massima inefficacia media percentuale per una azione di mitigazione che non sia
fortemente dannosa (es 2%)
  MOD_THRESHOLD = 100 % soglia di memorizzazione

% EOF
```

Vediamo ora il formato dei due ulteriori file di configurazione, come nel caso sopra, passando da un esempio.

```
% file specs.conf

% rischi possibili per il sistema (checklist)
<RISKS>
  % primo rischio: rottura del dispositivo di stampa
  12, % ID, si consiglia di partire da 1 e incrementare
  A65S, % risk_CL_code
  Tecnico, % category
  guasto del dispositivo di stampa, % description
  surriscaldamento | guasto elettrico, % causes
  aumento TTM | insoddisfazione cliente, % effects
  true; % il rischio ha un impatto relativo al valore del progetto
  % secondo rischio: fallimento del fornitore
  13, H72S ,Management, fallimento del fornitore, spesa eccessiva, mancanza di parti | aumento TTM, false;

% possibili azioni di mitigazione (checklist)
<MITIGATION>
  % prima azione, addestrare il personale
  1, % ID, si consiglia di partire da 1 e incrementare
  76354A, % action_CL_code
  addestrare il personale, % description
  false; % l'azione non ha un costo relativo al valore del progetto
  % seconda azione, potenziare il macchinario
  2, 14573H, potenziare il macchinario di stampa, true; % l'azione ha un costo relativo al valore del progetto

% possibili azioni di recovery (checklist)
<RECOVERY>
  % prima azione, risparmiare per riparare il dispositivo di stampa
  1, % ID, si consiglia di partire da 1 e incrementare
  99121R, % action_CL_code
  spese di riparazione del dispositivo di stampa; % description
  % seconda azione, cambiare tecnologia utilizzata per fare una certa cosa
  2, 13244S, spese di cambio di tecnologia;
  % terza azione, sostituire un macchinario
  7, 73336A, sostituzione macchinario;

% le tre sezioni non devono comparire necessariamente nell'ordine
% descritto sopra

% EOF
```

```
% file projects.conf

% ad ogni progetto si fa corrispondere una sezione <project>
% con alcune sottosezioni che descrivono la struttura del progetto.
% l'ordine con cui compaiono le sottosezioni non è importante, ma devono
% esserci tutte.
<PROJECT>
  [ATTRIBUTES]
     2, % revision, revisione del progetto
     closed, % state, stato del progetto
     % Inizio degli attributi 'Features'
     project pegasus, % name, nome del progetto
     3489f26z, % company ID, ID del progetto come identificato
            % all'interno della compagnia
     Giuseppe Mandragola, % costumer name, nome del cliente
     300000, % value, valore del progetto in euro
     Microchip Pentium III, % service_material, oggetto della fornitura
     10, % valore del primo attributo specificato precedentemente in init.conf
     Londra, % valore del secondo attributo specificato precedentemente in init.conf
     scarso, % valore del terzo attributo specificato precedentemente in init.conf
     20, % ...
     Sicilia,
     35,
     12000,
     US Robotics,
     buona,
     5,
     Silver;
  [RISKS]
     12,  % ID del rischio, uno di quelli specificati in specs.conf
     A2373H,  % company_risk_ID
     mitigation | acceptance,  % strategia utilizzata, una nell'insieme:
                        % transfer, mitigation, acceptance,
                        % mitigation | acceptance
     monitoring | recovery, % stato del rischio, uno nell'insieme:
                      % monitoring, mitigation, recovery,
                      % monitoring | recovery,
                      % mitigation | recovery,
                      % closed_verified,  closed_unverified.
     % in questo caso lo stato del rischio è monitoring + recovery.
     1,   % revisione in cui è stato introdotto
     20 | 18-2,  % storico impatto del rischio, in questo caso valori percentuale
     30 | 25-2,  % storico probabilità del rischio, valori percentuali
       % gli storici vanno dal valore iniziale a quello attuale,
       % secondo la direzione: vecchio -> nuovo.
       % i valori dopo il carattere '-' rappresentano l'ID dell'azione di mitigazione
       % che ha causato il cambiamento (se presente)
     guasto della testina della stampante, % description
     surriscaldamento, % causes
     insoddisfazione cliente; % effects
```

```
    % altri rischi associati a questo progetto ...
  [MIT_ACTS]
    2,  % ID dell'azione di mitigazione,
        % uno di quelli specificati in specs.conf
    30, % costo dell'azione di mitigazione,
         % può essere una percentuale oppure un
         % valore in euro (a seconda del valore di rel_act).
    potenziare il macchinario di stampa, % description
    1,  % revision number
    closed; % stato della azione di mitigazione
    % altre azioni di mitigazione associate a questo progetto ...
  [REC_ACTS]
    7,  % ID dell'azione di recovery,
        % uno di quelli specificati in specs.conf
    sostituzione parti guaste del macchinario; % description
    % altre azioni di recovery associate a questo progetto
  [MIT_BINDS] % descrive quali azioni di mitigazione
                % sono legate a quali rischi
    12, % ID del rischio, deve comparire nella sottosezione RISKS
    2,  % ID dell'azione di mitigazione,
        % deve comparire nella sottosezione MIT_ACTS
    6,  % Efficacia dell'azione in termini di impatto
        % del rischio (valore percentuale)
    -2; % Efficacia dell'azione in termini di probabilità
         % del rischio (valore percentuale, in questo caso l'aumenta)
  % altri legami tra rischi ed azioni di mitigazione
  [REC_BINDS] % descrive quali azioni di
                % mitigazione sono legate a quali rischi
    12, % ID del rischio, deve comparire nella sottosezione RISKS
    7,  % ID dell'azione di recovery, deve comparire
        % nella sottosezione REC_ACTS
    20000; % valore del campo risk_contingency
         % (assume un significato diverso a seconda di
         % rel_risk associato al rischio). Anche se
         % l'utente opera sempre in euro il
         % sistema si occuperà di convertire internamente
         % il campo risk_contingency dividendo per il valore
         % del progetto qualora rel_risk per il rischio sia a
         % valore true.
  % altri legami tra rischi ed azioni di recovery

% ULTERIORI PROJECT
% ...

% EOF
```

3 Base di Dati

Si è deciso di realizzare la base di dati del sistema (come già accennato) utilizzando un database MySQL, in questa sezione vediamo la struttura dei dati memorizzati in memoria di massa.

Nella precedente sezione è stato illustrato come un progetto, una specifica di rischio o una specifica di azione di mitigazione/recovery vengono rappresentate in memoria, non si può pensare però di tenere in memoria enormi quantità di questi dati. È inoltre necessario che i dati siano disponibili anche una volta che il sistema è stato riavviato per questioni di robustezza. Ci si affida allora ad un database MySQL per l'immagazzinamento e l'accesso efficiente ai dati di cui abbiamo bisogno.

3.1 Tabelle nel database

Vediamo la struttura delle tabelle, ricordando che la base di dati è di tipo relazionale.

Tabelle relative ad aspetti generali del funzionamento del sistema

- **parameters(<u>par_name</u>, value)**
  Questa tabella si occupa di memorizzare il valore di tutti i parametri di interesse per l'inferenza, ad esempio il valore di K.

- **attributes(<u>att_name</u>, group, category, type, additional_info)**
  In questa tabella si memorizzano tutti gli attributi specificati dall'utente per i progetti (recuperati dal file init.conf) e per ognuno di essi si tiene traccia del gruppo a cui appartiene e della categoria di cui fa parte (come descritto dall'utente stesso nel medesimo file). La categoria di un attributo può essere 1 oppure 2 (1 = relativo al livello del rischio, 2 = relativo all'impatto strategico). Inoltre si tiene traccia del tipo dell'attributo (uno tra enumerato E, numerico N, percentuale P, stringa S oppure custom C). Ricordiamo che per gli attributi enumerati si devono specificare nel file di configurazione i diversi possibili valori, mentre per gli attributi numerici si può specificare un'unità di misura. Il campo additional_info è allora un campo che contiene una stringa che descrive i possibili valori di un attributo enumerato (separati da |, es: "buono|medio|cattivo"), oppure l'unità di misura per un attributo numerico (ad esempio "km"). Per gli attributi di tipo custom il campo additional_info conterrà i possibili valori per l'attributo e subito a seguire la matrice di somiglianza scritta riga per riga (ad esempio "Sicilia|Sardegna|Corsica [100 70 20 70 100 30 20 30 100]").
  Notare che per gli attributi percentuali e stringhe il campo additional_info risulta inutilizzato, ma questo non è un problema in quanto comunque questa tabella avrà dimensioni abbastanza contenute.

Tabelle relative alle specifiche dei rischi, delle azioni di mitigazione e delle azioni di recovery

- **risks(<u>risk_ID</u>, risk_CL_code, category, description, causes, effects, rel_risk)**
  Questa tabella memorizza tutte le specifiche dei rischi, queste si ottengono dal parsing del file specs.conf. Come si vede l'ID del rischio deve essere univoco. I campi sono omonimi a quelli presenti nella classe Risk e dunque il loro significato è noto. Gli attributi causes e effects contengono liste di frasi globalmente viste come stringhe, queste liste sono rappresentate come stringhe dove le frasi sono separate da un separatore '|'.

- **mitigation_actions(<u>mitigation_ID</u>, action_CL_code, rel_act, description)**
  Questa tabella memorizza tutte le specifiche di azioni di mitigazione, queste si ottengono dal parsing del file specs.conf. Come si vede l'ID dell'azione di mitigazione deve essere univoco. I campi sono omonimi a quelli presenti nella classe Mit_Action e dunque il loro significato è noto.

- **recovery_actions(<u>recovery_ID</u>, action_CL_code, description)**
  Questa tabella memorizza tutte le specifiche di azioni di recovery, queste si ottengono dal parsing del file specs.conf. Come si vede l'ID dell'azione di recovery deve essere univoco. I campi sono omonimi a quelli presenti nella classe Rec_Action e dunque il loro significato è noto.

Tabelle relative ai vari aspetti dei progetti memorizzati, iniziano tutte con "proj_"

- **proj_attributes(<u>project_ID</u>, revision, state, project_name, company_ID, customer_name, value, service_material, <additional attributes>, use_for_inference)**
  Questa tabella memorizza tutti gli attributi per ogni progetto, notare che in essa si trovano non solo gli attributi predefiniti, ma anche quelli definiti dall'utente (indicati con **<additional attributes>**). L'unica chiave in questa tabella è data dal project_ID. Gli attributi definiti dall'utente vengono memorizzati nell'ordine che l'utente ci ha fornito nel file di configurazione init.conf. Il campo use_for_inference, inizialmente a true per tutti i progetti inseriti dall'utente, viene posto a false quando il progetto in chiusura viene ritenuto sufficientemente simile a quelli presenti nella base di conoscenza e dunque inutile per l'inferenza (vedere la sezione 4).

- **proj_risks(<u>project_ID, risk_ID</u>, company_risk_ID, strategy, state, revision, impact, probability, description, causes, effects)**
  Questa tabella tiene traccia di tutti i rischi associati ad un certo progetto, notare infatti che la chiave primaria è in questo caso costituita dalla coppia (project_ID, risk_ID). Gli altri campi sono omonimi a quelli che si trovano nella classe Proj_Risk, e dunque il loro significato è noto. Per quanto riguarda i campi impact e probability, questi non sono numeri ma bensì liste realizzate come stringhe di coppie (elemento numerico, mitigation_ID) separati da separatori '|'. Gli attributi causes e effects contengono liste di frasi globalmente intese come stringhe, queste liste sono rappresentate come stringhe dove le frasi sono separate da un separatore '|'.

- **proj_mit_actions(<u>project_ID, mitigation_ID</u>, cost, description, revision, state)**
  Questa tabella tiene traccia di tutte le azioni di mitigazione previste per un certo progetto, notare infatti che come chiave primaria si utilizza la coppia project_ID, mitigation_ID. I campi sono gli stessi che si trovano nella classe Proj_Mit_Action e dunque non hanno bisogno di essere spiegati.

- **proj_rec_actions(<u>project_ID, recovery_ID</u>, description)**
  Questa tabella tiene traccia di tutte le azioni di recovery previste per un certo progetto, notare infatti che come chiave primaria si utilizza la coppia project_ID, recovery_ID. I campi sono gli stessi che si trovano nella classe Proj_Rec_Action e dunque non hanno bisogno di essere spiegati.

- **proj_mit_binds(<u>project_ID, risk_ID, mitigation_ID</u>, impact_eff, probability_eff)**
  Questa tabella fornisce tutte le associazioni tra rischi all'interno dei progetti e le azioni di mitigazione all'interno del medesimo progetto. Visto che infatti la chiave primaria è data dalla terna project_ID, risk_ID, mitigation_ID è ovvio che ogni tupla identifica univocamente un rischio e una azione di mitigazione all'interno di un certo progetto. Gli altri campi corrispondono a impact_effectiveness e probability_effectiveness viste esaminando la struttura della classe Proj_Mit_Bind.

- **proj_rec_binds(<u>project_ID, risk_ID</u>, recovery_ID, risk_contingency, original_RI)**
  Questa tabella fornisce tutte le associazioni tra rischi all'interno dei progetti e le azioni di recovery all'interno del medesimo progetto. Ricordiamo che per ogni rischio all'interno di un progetto ci può essere al più una unica azione di recovery e quindi è giusto considerare come chiave primaria la coppia project_ID, risk_ID. Se ad un rischio in un determinato non è associata nessuna azione di recovery allora semplicemente non avremo una tupla in questa tabella relativa a tale rischio in quel progetto. Gli altri campi corrispondono al risk_contingency e original_ID visti esaminando la struttura della classe rischio. Notare che per le azioni di recovery aggiunte direttamente in fase di inferenza il campo original_RI contiene sempre i valori iniziali per probabilità e impatto del rischio moltiplicati tra loro come descritto in sezione 2, dunque non sarebbe necessario memorizzarlo a parte. Per semplicità si è però deciso memorizzare comunque questa informazione, che risulta molto comoda nell'estrazione dei dati dal database. Questo introduce una minima ridondanza nella base di dati, ma non è un grosso problema.

- **proj_copies(<u>project_ID</u>, <u>copy_ID</u>)**
  Questa tabella fornisce per ogni progetto in stato attivo (non chiuso) il project_ID del progetto copia che è stato creato al termine dell'inferenza subito prima della fase di validazione (vedere la sezione 4). Il progetto copia serve per valutare (alla chiusura del progetto effettivo) quanto l'utente ha modificato il risultato proposto dal sistema. Quando un progetto non copia diventa chiuso, la tupla ad esso associata in questa tabella viene eliminata in quanto il progetto copia non è più necessario e viene eliminato dalla base di dati. È ovvio che i progetti il cui project_ID è contenuto in una tupla di questa tabella nel campo copy_ID sono delle copie e pertanto *non vanno considerati nel processo di inferenza.* Anche l'attributo copy_ID è chiave perché notiamo che un progetto copia può esser copia di un solo progetto effettivo.

3.2 Vincoli di integrità

Vi sono alcuni vincoli di integrità referenziale da considerare per il corretto funzionamento della base di dati. Il rispetto di questi vincoli è garantito dal Database Management System (DBMS) e sono tutti vincoli di chiave esterna.

<u>Tabella proj_risks</u>

1. Il project_ID presente nella tabella proj_risks deve essere anche presente nella tabella proj_attributes: *l'attributo project_ID nella tabella proj_risks è chiave esterna referenziante la tabella proj_attributes*.
2. Il risk_ID presente nella tabella proj_risks deve essere anche presente nella tabella risks: *l'attributo risk_ID nella tabella proj_risks è chiave esterna referenziante la tabella risks.*

Tabella proj_mit_actions

3. Il project_ID presente nella tabella proj_mit_actions deve essere anche presente nella tabella proj_attributes: *l'attributo project_ID nella tabella proj_mit_actions è chiave esterna referenziante la tabella proj_attributes*.
4. Il mitigation_ID presente nella tabella proj_mit_actions deve essere anche presente nella tabella mitigation_actions: *l'attributo mitigation_ID nella tabella proj_mit_actions è chiave esterna referenziante la tabella mitigation_actions.*

Tabella proj_rec_actions

5. Il project_ID presente nella tabella proj_rec_actions deve essere anche presente nella tabella proj_attributes: *l'attributo project_ID nella tabella proj_rec_actions è chiave esterna referenziante la tabella proj_attributes*.
6. Il recovery_ID presente nella tabella proj_rec_actions deve essere anche presente nella tabella recovery_actions: *l'attributo recovery_ID nella tabella proj_rec_actions è chiave esterna referenziante la tabella recovery_actions.*

Tabella proj_mit_binds

7. La coppia (project_ID, risk_ID) presente nella tabella proj_mit_binds deve essere anche presente nella tabella proj_risks: *la coppia project_ID, risk_ID nella tabella proj_mit_binds è chiave esterna referenziante la tabella proj_risks*.
8. La coppia (project_ID, mitigation_ID) presente nella tabella proj_mit_binds deve essere anche presente nella tabella proj_mit_actions: *la coppia project_ID, mitigation_ID nella tabella proj_mit_binds è chiave esterna referenziante la tabella project_mit_actions.*

Tabella proj_rec_binds

9. La coppia (project_ID, risk_ID) presente nella tabella proj_rec_binds deve essere anche presente nella tabella proj_risks: *la coppia project_ID, risk_ID nella tabella proj_rec_binds è chiave esterna referenziante la tabella proj_risks*.
10. La coppia (project_ID, recovery_ID) presente nella tabella proj_rec_binds deve essere anche presente nella tabella proj_rec_actions: *la coppia project_ID, recovery_ID nella tabella proj_rec_binds è chiave esterna referenziante la tabella project_rec_actions.*

Tabella proj_copies

11. Il project_ID presente nella tabella proj_copies deve essere anche presente nella tabella proj_attributes: *l'attributo project_ID nella tabella proj_copies è chiave esterna referenziante la tabella proj_attributes*.
12. Il copy_ID presente nella tabella proj_copies deve essere anche presente nella tabella proj_attributes: *l'attributo copy_ID nella tabella proj_copies è chiave esterna referenziante la tabella proj_attributes (attributo project_ID)*.

Quando uno qualsiasi di questi vincoli viene violato il DBMS impedisce l'operazione restituendo errore, ossia si applica ***RESTRICT*** (sia su delete che su update). Questo significa che le operazioni di inserimento e rimozione andranno fatte in un certo ordine.

4 Modulo di Inferenza

In questa sezione diamo i dettagli del funzionamento del modulo di inferenza, e del processo di inferenza stesso.

Il modulo di inferenza ha due compiti fondamentali:

1. **Case Based Reasoning sui nuovi progetti inseriti dall'utente**
   Dato un nuovo progetto di cui sono noti gli attributi, il sistema fornisce i rischi significativi che affliggono il nuovo progetto con i relativi parametri, le azioni di mitigazione consigliate con la relativa efficacia prospettica e costo, e le azioni di recovery raccomandate.
2. **Gestione della chiusura di un progetto**
   Prima della chiusura di un progetto attivo, esso costituisce sempre un caso utile al processo di CBR di cui sopra. Quando invece un progetto si chiude, dobbiamo decidere se effettivamente il progetto risulta sufficientemente diverso dagli altri nella base di conoscenza da essere memorizzato definitivamente come caso utile al CBR.

Come già detto, il modulo di inferenza deve interagire con il modulo di gestione della conoscenza: si prevede dunque che esso abbia la facoltà di invocare delle operazioni offerte come servizi dal modulo di gestione della conoscenza. Non entreremo nel dettaglio dell'interfaccia di queste operazioni, che risulta di interesse puramente implementativo.

4.1 Case Based Reasoning (CBR)

La funzione di case based reasoning del modulo di inferenza entra in gioco solamente quando viene presentato un nuovo progetto. Una volta che il processo di inferenza è terminato su tale progetto, si delega al modulo di interfaccia con il cliente la gestione delle modifiche post-inferenza al progetto (ovviamente attraverso il modulo di gestione della conoscenza).

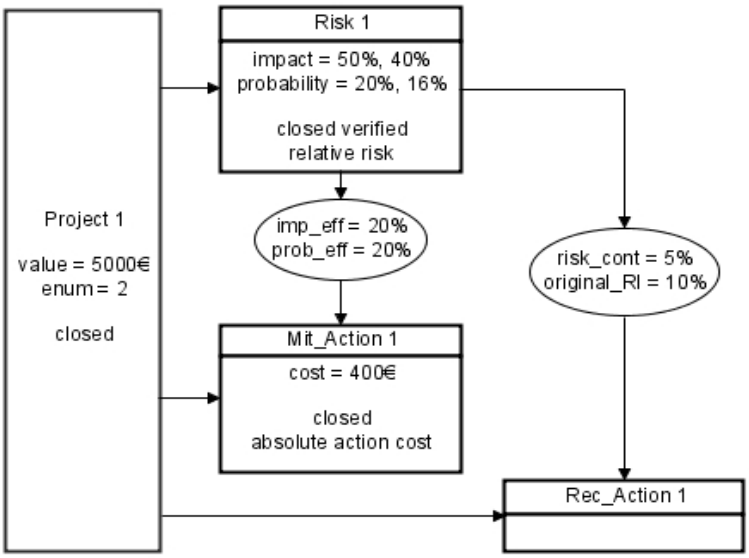
Vediamo allora il funzionamento di ogni singola fase del processo di inferenza, consci del fatto che ogni operazione fatta sui progetti viene opportunamente supportata dal modulo di gestione della conoscenza.

Per semplificare la comprensione del lettore delle 5 fasi del processo di ragionamento basato sui casi, vedremo un esempio completo di inferenza prendendo come riferimento una semplice base di conoscenza. Ogni volta che in questo documento verranno mostrati dati relativi al semplice esempio in esame, useremo una sezione ben delimitata come la seguente. Tutti i dati relativi a questo semplice esempio sono stati accuratamente verificati usando l'implementazione proposta per il CBR Risk Management System

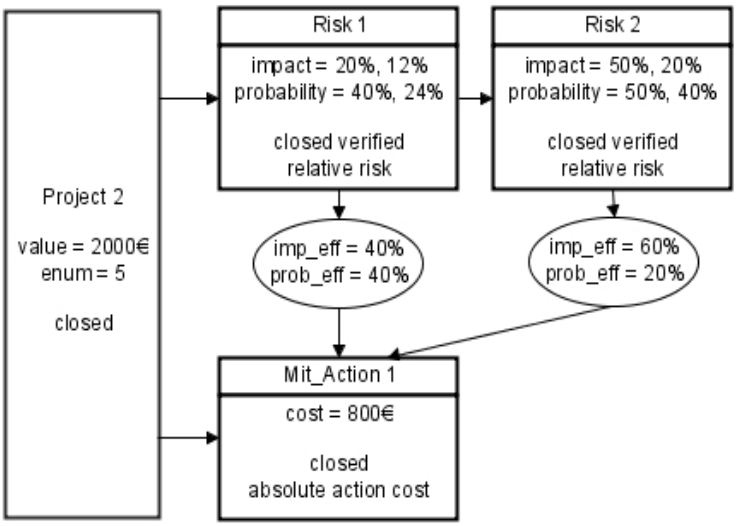
ESEMPIO

L'esempio di inferenza che prendiamo in considerazione prevede la situazione seguente

- L'unico attributo aggiuntivo rispetto a quelli predefiniti nel sistema (configurato dall'utente) è un attributo di tipo enumerato [E] di nome enum, il cui valore appartiene all'insieme dei seguenti 5 possibili valori: $\{1,2,3,4,5\}$.
- La base di conoscenza contiene solamente due casi passati, che possono venir descritti in modo intuitivo tramite le seguenti figure.
  - **Progetto 1**



  - **Progetto 2**



- Il nuovo progetto fornito dall'utente ha le seguenti specifiche (sono riportate solo le informazioni) utili alla comprensione del processo di inferenza, rappresentate tramite un'immagine.
  - **Nuovo Progetto**

Fase 1: Estrazione dei progetti simili

L'estrazione dei progetti simili si basa sul calcolo di un valore di distanza del nuovo progetto da tutti gli altri progetti all'interno della base di conoscenza (casi passati). La distanza viene calcolata facendo riferimento solamente alle Caratteristiche (Features) dei progetti.
Quando l'utente inoltra un progetto al modulo di inferenza, questo ha già un certo valore in ognuna delle caratteristiche del progetto. Da ora in poi i termini "caratteristiche", "parametri", "attributi" e "campi" sono utilizzati in modo intercambiabile e fanno sempre riferimento ad elementi all'interno del campo "Features" di un progetto.

Gli attributi di un progetto sono sempre suddivisi in gruppi configurati alla prima accensione del sistema tramite il file init.conf, caratteristiche appartenenti ad uno stesso gruppo descrivono uno specifico aspetto del progetto a seconda dell'esperienza dell'esperto umano che ha configurato il sistema in prima istanza.

Indichiamo con $\{G_j\}$ l'insieme dei gruppi.

Ad ogni gruppo appartiene un sottoinsieme degli attributi dei progetti, il processo di inferenza considera diversi gruppi come *diversi aspetti dei progetti da usare per la somiglianza*.

Quando l'utente ha terminato la specifica del nuovo progetto e si deve procedere alla valutazione intelligente dei rischi e delle azioni di mitigazione, viene presentata un'interfaccia dotata di opportune checkbox che consente di scegliere un sottoinsieme di attributi da usare per l'estrazione dei progetti simili.
L'utente potrebbe scegliere un sottoinsieme di attributi che contiene attributi appartenenti a tutti i gruppi, oppure può anche essere il caso in cui non vi sono attributi appartenenti ad un particolare gruppo $\bar{G}$ nel sottoinsieme scelto.

Per ogni gruppo di cui l'utente ha selezionato almeno un attributo, si va a calcolare la distanza di tutti i progetti nella base di conoscenza dal nuovo progetto (usando gli attributi specificati dall'utente per il gruppo).

Siano gli attributi dei progetti gli elementi dell'insieme $\{Att_k\}$.

Si scorrono allora tutti i progetti nella base di conoscenza calcolando in ogni progetto per ogni gruppo di cui è stato selezionato almeno un attributo un <u>grado di somiglianza non normalizzato</u>, definito GS_{ij}^{nn} per l'i-esimo progetto relativo al j-esimo gruppo.

È anche possibile configurare per un determinato attributo selezionato un <u>peso</u> AW_k^e (Attribute Weight, extraction) diverso dal valore di default (1) che dia più importanza o meno importanza all'attributo relativamente agli altri attributi scelti, questo viene realizzato tramite un opportuno campo di testo modificabile dall'utente a fianco delle checkbox.

Ci fa comodo, dato l'insieme degli $\{AW_k^e\}$ dei pesi definiti dall'utente o meno (valori di default), usarlo per definire un altro insieme $\{AWj_k^e\}$ dei pesi nell'ottica del gruppo G_j, dove:

$$AWj_k^e = \begin{cases} AW_k^e & se\ Att_k\ appartiene\ a\ G_j \\ 0 & se\ Att_k\ non\ appartiene\ a\ G_j \end{cases}$$

Per un attributo non selezionato per l'estrazione si avrà $AW_k^e = 0 \rightarrow AWj_k^e = 0$.

Il peso di ogni attributo nell'ottica del gruppo G_j viene utilizzato per calcolare GS_{ij}^{nn}.
Quando allora stiamo esaminando un particolare progetto (l'i-esimo) per valutare GS_{ij}^{nn}, si scorrono tutti gli attributi scelti dall'utente appartenenti a G_j e si procede come segue.

Chiamiamo A_k^n il valore dell'attributo nel nuovo progetto (new) fornito dall'utente, A_k^i il valore dell'attributo nel progetto i-esimo in considerazione e infine Att_k il nome dell'attributo.
Si calcola per ogni attributo selezionato un particolare valore detto livello di somiglianza, compreso tra 0 ed 1, che si va a sommare (moltiplicato per il peso AWj_k^e) al totale per GS_{ij}^{nn}. Il calcolo del livello di somiglianza per l'attributo Att_k nel progetto i-esimo (LS_k^i) è diverso a seconda del tipo di attributo, in particolare ci fa comodo definire un livello di diversità come 1-livello di somiglianza ($LD_k^i = 1 - LS_k^i$).

È ovvio allora che GS_{ij}^{nn}, calcolato nel progetto i-esimo per il gruppo G_j, non andrà semplicemente da 0 al numero di attributi selezionati per il gruppo G_j, ma spazierà nell'intervallo $[0, \sum_k AWj_k^e]$.

Vediamo allora come ci si comporta con i diversi tipi di attributi:

- *Attributo di tipo stringa [S]*
  In questo caso si fa un matching case sensitive per vedere se la parola o la frase memorizzata è la stessa o meno:
  - $LS_k^i = 1$ se la stringa è la stessa ($A_k^n = A_k^i$)
  - $LS_k^i = 0$ altrimenti ($A_k^n \neq A_k^i$)
- *Attributo di tipo enumerato [E]*
  In questo caso è noto il numero dei possibili valori per l'enumerato, sia questo numero pari a N_{val}. Per questo risulta particolarmente semplice calcolare il livello di diversità come segue.
  $$LD_k^i = \frac{|A_k^n - A_k^i|}{N_{val} - 1}$$
  Notare infatti che la differenza a numeratore non è una differenza tra stringhe, ma una differenza numerica in quanto gli enumerati sono visti dal sistema come valori numerici naturali (a partire da 0, crescendo).
  Per cui si da infine:
  $$LS_k^i = 1 - LD_k^i$$

- *Attributo di tipo percentuale [P]*

  In questo caso è noto il fatto che l'attributo è riferito al valore 100% e dunque è possibile fare un'operazione simile a quella vista per gli attributi enumerati, come segue:

  $$LD_k^i = \frac{|A_k^n - A_k^i|}{100}$$

  e dunque:

  $$LS_k^i = 1 - LD_k^i$$

- *Attributo di tipo custom [C]*

  In questo caso il valore del livello di somiglianza si legge direttamente dalla matrice di somiglianza, tenendo presente che i valori scritti nella matrice di somiglianza sono numeri reali tra 0 e 100 che rappresentano delle percentuali.

  Se allora A_k^i è il t-esimo valore specificato dall'utente per l'attributo Att_k in init.conf e A_k^n è il q-esimo valore allora si impone semplicemente:

  $$LS_k^i = \frac{M_{tq}}{100}$$

  Dove M è la matrice di somiglianza associata all'attributo Att_k.

- *Attributo di tipo numerico [N]*

  In questo caso non possiamo dire a priori quale è il range di variabilità dell'attributo e non è facile trovare un valore a cui riferire la distanza tra A_k^n e A_k^i.

  Il sistema è in grado di fornire un livello di somiglianza senza la necessità che l'utente intervenga specificando un range per ogni attributo, tuttavia si vuole dare anche all'utente la capacità di intervenire qualora il funzionamento di default stabilito per il sistema non sia ritenuto idoneo.

  Questo intervento dell'utente è fatto tramite interfaccia grafica direttamente durante la specifica degli attributi per l'inferenza.

  L'algoritmo di calcolo del livello di somiglianza è il seguente:

  $$D = A_k^i - A_k^n$$

  $$LD_k^i = \begin{cases} \frac{D}{R_1} & if\ D > 0 \\ \frac{|D|}{R_2} & if\ D < 0 \\ 0 & if\ D = 0 \end{cases}$$

  $$LS_k^i = 1 - LD_k^i \ \ (0\ minimo)$$

  Di default si ha che:

  $$\begin{cases} R_1 = \frac{A_k^n}{2} \\ R_2 = \frac{A_k^n}{2} \end{cases}$$

  La capacità dell'utente di modificare questo processo si basa sulla possibilità di modificare R_1 ed R_2 dal loro valore di default. In particolare, quando l'utente seleziona una checkbox che

corrisponde ad un attributo numerico, gli si presenta una frase con due campi di testo modificabili (in blu o in rosso), vediamo degli esempi:

- l'utente seleziona un attributo con valore nel nuovo progetto pari a 50000.
  compare:
  "ex.
  [25000] is smaller and far
  [75000] is greater and far "

- l'utente seleziona un attributo a con valore nel nuovo progetto pari a 10.
  compare:
  "ex.
  [5] is smaller and far
  [15] is bigger and far "
- l'utente seleziona un attributo a con valore nel nuovo progetto pari a 0.
  compare:
  "ex.
  [no value] is smaller and far
  [no value] is bigger and far
  change the red values or uncheck the box to proceed"

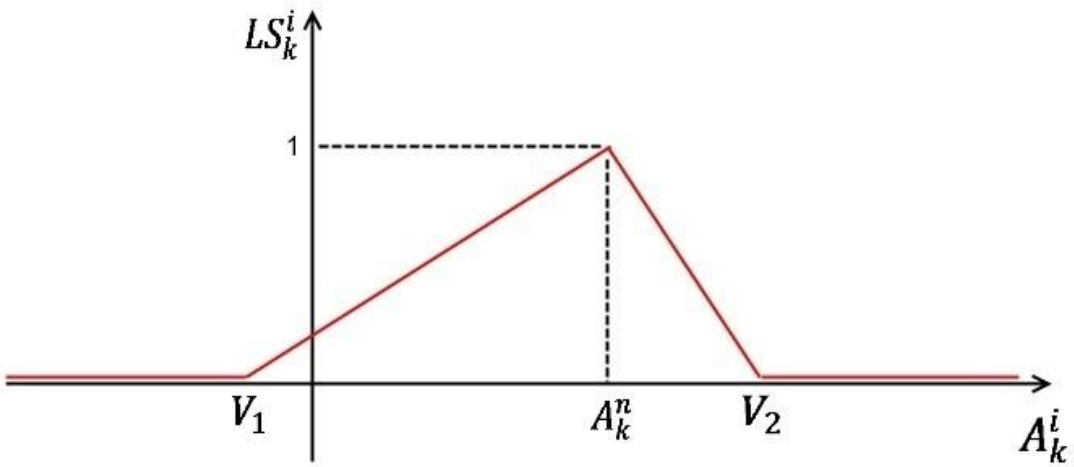
A questo punto l'utente può modificare i valori in blu nelle caselle di testo e deve modificarli qualora il valore per il nuovo progetto sia 0 (in tal caso il valore da modificare è mostrato in rosso), o deselezionare la checkbox.

Definendo V_1 il valore della prima casella di testo e V_2 quello della seconda allora si calcola:

$$\begin{cases} R_1 = V_2 - A_k^n \\ R_2 = A_k^n - V_1 \end{cases}$$

Per cui è ovvio che se l'utente non modifica nulla il comportamento rimane quello standard con R_1 ed R_2 pari alla metà del valore dell'attributo nel nuovo progetto.

In pratica allora l'andamento del livello di somiglianza in funzione del valore dell'attributo nel progetto estratto (A_k^i) è il seguente:



A questo punto è facile capire come alla fine si calcola per ogni progetto e relativamente ad ogni gruppo un grado di somiglianza non normalizzato, come segue:

$$GS_{ij}^{nn} = \sum_{k} \left(LS_k^i * AWj_k^e \right) \qquad \begin{array}{c} Grado\ di\ somiglianza\ non\ normalizzato \\ del\ progetto\ i-esimo\ nel\ gruppo\ j-esimo \end{array}$$

Per ogni gruppo al termine si scelgono i primi MAX_PROJ_PER_GROUP con grado di somiglianza non normalizzato più alto (dove MAX_PROJ_PER_GROUP è un parametro di inferenza configurabile dall'utente in prima istanza tramite il file init.conf e poi tramite GUI).

Al termine della procedura di estrazione dei progetti il sistema ha a disposizione un insieme di alcuni gruppi in ognuno dei quali sono presenti dei progetti più o meno simili a quello fornito dall'utente, la situazione è simile a quella riportata nella seguente tabella, dove per ipotesi si è assunto MAX_PROJ_PER_GROUP = 6.

G_1	G_2	G_3	G_4	G_5
P_1	P_3	P_7	P_2	P_{10}
P_3	P_7	P_3		P_{11}
P_5	P_4	P_4		P_{12}
	P_2	P_5		
	P_{10}	P_6		
		P_1		

Questa costituisce la cosiddetta <u>tabella dei progetti estratti</u>.

ESEMPIO

Per quanto riguarda l'esempio che abbiamo preso in considerazione, facciamo le seguenti ipotesi:

- Tutti gli attributi dei progetti fanno parte del medesimo gruppo, il gruppo G_1
- L'utente ha selezionato per l'estrazione i due attributi enum e value
- Il peso specificato dall'utente è unitario per entrambi gli attributi (valore di default)
- Per l'attributo value l'utente mantiene i valori di default per R_1 ed R_2, in particolare:
  $R_1 = R_2 = 1250$

Nuovo progetto → value = 2500 ed enum = 4
Applicando quanto appena visto:

- Progetto 1
  value = 5000 ed enum = 2
  $GS_{11}^{nn} = 0.5\ (enum) + \ 0.0(value) = 0.5$

- Progetto 2
  value = 2000 ed enum = 5
  $GS_{21}^{nn} = 0.75\ (enum) + \ 0.6(value) = 1.35$

Il risultato della fase 1 risulta allora:

Gruppo G_1
Progetto 2
Progetto 1

Fase 2: Ranking dei progetti estratti

Il processo di calcolo del grado di somiglianza non normalizzato come descritto nella Fase 1 è valido per capire come funziona all'interno del sistema il processo di estrazione dei progetti simili, ma in realtà bisogna entrare un po' più nel dettaglio per capire come si procede al ranking dei progetti estratti.

Al momento in cui all'utente è richiesto di selezionare gli attributi per l'estrazione dei progetti, quest'ultimo può indicare anche un insieme di attributi utilizzati per il ranking. In particolare, all'utente si presenta nella stessa interfaccia una lista di attributi da cui può scegliere (tramite checkbox) quelli per effettuare l'estrazione dei progetti simili e una lista di attributi da cui può invece scegliere (sempre tramite checkbox) ulteriori attributi (anche parzialmente sovrapposti agli attributi precedenti) che verranno utilizzati per effettuare il ranking dei progetti una volta che essi sono stati estratti.
Anche il campo che consente di specificare il peso associato ad un attributo è duplicato e dunque è possibile associare ad un attributo un peso diverso per l'estrazione dei progetti simili e per il ranking (AW_k^r), anche in questo caso un attributo non selezionato per il ranking avrà $AW_k^r = 0$. È inoltre presente un controllo che consente all'utente di scegliere tra

- Ranking basato solo sugli attributi scelti per il ranking
- Ranking basato sia sugli attributi di ranking che sugli attributi usati per l'estrazione (*extraction based ranking*)

La seconda opzione è fornita all'utente perché consente di fare un ranking dei progetti senza per forza selezionare gli attributi per il ranking (volendo usare gli stessi dell'estrazione) e perché consente di differenziare uno stesso progetto in diversi gruppi (vedere oltre).
Chiamiamo allora $\{Att_{kj}^e\}$ l'insieme degli attributi (Att) scelti dall'utente per l'estrazione (e) appartenenti al gruppo G_j. Chiamiamo invece $\{Att_k^r\}$ l'insieme degli attributi (Att) scelti dall'utente per il ranking (r). Sia inoltre LS_k^i il livello di somiglianza calcolato per l'attributo k-esimo nel progetto i-esimo come definito precedentemente.

Quando si scorrono i progetti presenti nella base di conoscenza si calcola allora per ogni gruppo di attributi non vuoto due gradi di somiglianza in ogni progetto:

- il grado di somiglianza non normalizzato visto precedentemente, GS_{ij}^{nn}.
- il <u>grado di somiglianza non normalizzato per il ranking</u> (che chiamiamo $GS_{ij}^{nn\ rank}$), questo grado di somiglianza si calcola come fatto per il grado di somiglianza non normalizzato ma considerando come insieme di attributi solo gli attributi per il ranking oppure attributi per il ranking e attributi per l'estrazione (a seconda dell'opzione scelta dall'utente), dunque:
  - *ranking basato solo sugli attributi per il ranking*

    $$GS_{ij}^{nn\ rank} = \sum_k \left(LS_k^i * AW_k^r\right) \quad \begin{matrix} Grado\ di\ somiglianza\ non\ normalizzato \\ per\ il\ ranking\ del\ progetto\ i-esimo \\ nel\ gruppo\ j-esimo \end{matrix}$$

    Notare che non si ha dipendenza da j e dunque il $GS_{ij}^{nn\ rank}$ risulta lo stesso per ogni gruppo (come è giusto che sia in questo caso).
  - *ranking basato sugli attributi per il ranking e sugli attributi per l'estrazione*

    $$GS_{ij}^{nn\ rank} = \sum_k \left(LS_k^i * AWj_k^e\right) + \sum_k \left(LS_k^i * AW_k^r\right)$$

    $$\begin{matrix} Grado\ di\ somiglianza\ non\ normalizzato \\ per\ il\ ranking\ del\ progetto\ i-esimo \\ nel\ gruppo\ j-esimo \end{matrix}$$

    Stavolta si ha dipendenza da j in quanto il peso di un attributo per quanto riguarda l'estrazione è considerato essere nullo se l'attributo non appartiene al gruppo. In questo secondo caso notiamo che:

    $$GS_{ij}^{nn\ rank} = GS_{ij}^{nn} + \sum_k \left(LS_k^i * AW_k^r\right) \geq GS_{ij}^{nn}$$

Alla fine dell'estrazione (fase 1) allora abbiamo una struttura tabellare come mostrato precedentemente dove sono stati isolati per ogni gruppo i MAX_PROJ_PER_GROUP progetti con grado di somiglianza non normalizzato (per l'estrazione) più alto.

Questi progetti però non hanno ancora subito un vero ranking e non abbiamo ancora un ordinamento che ci dice effettivamente quanto ogni progetto sia somigliante nell'ambito di ogni gruppo. In realtà l'informazione necessaria per ordinare i progetti in base all'effettiva somiglianza (e l'unico parametro effettivamente da memorizzare per le successive fasi dell'inferenza) è proprio $GS_{ij}^{nn\ rank}$.

In realtà il rank di un progetto nell'ambito di un gruppo che viene memorizzato prende il nome di <u>*grado di somiglianza*</u> GS_{ij} (inteso come grado di somiglianza per il ranking e *normalizzato*) e si calcola come segue:

- Se si è fatto un ranking basato solo sugli attributi per il ranking

  $$GS_{ij} = \frac{GS_{ij}^{nn\ rank}}{\sum_k AW_k^r} \quad \begin{matrix} Grado\ di\ somiglianza\ del\ progetto\ i-esimo \\ nel\ gruppo\ j-esimo(memorizzato\ per\ le \\ successive\ fasi\ dell'inferenza) \end{matrix}$$

  Il quale dunque risulta indipendente da j.

- Se si è fatto un ranking basato sia sugli attributi per il ranking che sugli attributi per l'estrazione (extraction based ranking):

$$GS_{ij} = \frac{GS_{ij}^{nn\ rank}}{\sum_k AW_k^r + \sum_k AWj_k^e} \qquad \begin{array}{c} Grado\ di\ somiglianza\ del\ progetto\ i-esimo \\ nel\ gruppo\ j-esimo(memorizzato\ per\ le \\ successive\ fasi\ dell'inferenza) \end{array}$$

GS_{ij} è dunque un numero compreso tra 0 ed 1 e viene mantenuto per ogni progetto presente nella tabella risultato della fase 1 nell'ambito di un particolare gruppo. Allo stesso progetto possono quindi venire associati molteplici GS_{ij} in quanto può comparire diverse volte nella struttura tabellare, in diversi gruppi.

La forma finale della <u>*tabella dei progetti estratti*</u> è allora la seguente:

G_1	G_2	G_3	G_4	G_5
P_1, GS_{11}	P_3, GS_{32}	P_7, GS_{73}	P_2, GS_{24}	P_{10}, GS_{105}
P_3, GS_{31}	P_7, GS_{72}	P_3, GS_{33}		P_{11}, GS_{115}
P_5, GS_{51}	P_4, GS_{42}	P_4, GS_{43}		P_{12}, GS_{125}
	P_2, GS_{22}	P_5, GS_{53}		
	P_{10}, GS_{102}	P_6, GS_{63}		
		P_1, GS_{13}		

Come risultato di questa fase, all'interno della tabella dei progetti estratti vista al termine della fase 1 sono stato inserite delle nuove informazioni: adesso ad ogni progetto nell'ambito di un particolare gruppo è associato un grado di somiglianza GS_{ij} (calcolato con le formule appena viste), il quale è un valore tra 0 e 1 indicativo della somiglianza del progetto a cui è associato (i-esimo) con il nuovo progetto fornito dall'utente nell'ambito del gruppo G_j.

È interessante notare che qualora si riesca a trovare un progetto avente grado di somiglianza nell'ambito di un particolare gruppo pari a $GS_{ij} = 1$ allora significa che l'i-esimo progetto estratto corrisponde perfettamente al nuovo progetto nell'ambito del gruppo j-esimo. Questo tuttavia *non* implica che si possano utilizzare direttamente i rischi e le azioni presenti nel progetto i-esimo senza adattamento al nuovo caso.

Questo deriva da due diverse considerazioni:

- Non è assolutamente detto che per il ranking si siano considerati tutti gli attributi dei progetti, quindi se un progetto risulta avere $GS_{ij} = 1$ non è detto che sia identico al nuovo progetto considerando tutti gli attributi
- GS_{ij} è relativo al gruppo j-esimo: se anche scegliessero tutti gli attributi combinando estrazione e ranking potrebbe comunque succedere che il progetto i-esimo risulti identico al nuovo progetto nell'ambito del gruppo j-esimo, ma diverso in altri gruppi

Dato il modo in cui si recuperano e valutano (ordinano) i casi passati, in questo tipo di sistema CBR dovrà sempre esserci una fase di adattamento delle soluzioni recuperate utilizzando una molteplicità di casi estratti che risultano simili al nuovo progetto sotto vari punti di vista (gruppi).

ESEMPIO

Per quanto il semplice esempio di inferenza in esame, facciamo le seguenti ipotesi:

- L'utente non seleziona alcun nuovo attributo per il ranking
- L'utente ha scelto di usare l'opzione di extraction based ranking

Applicando quanto appena visto:

- Progetto 1

$$GS_{11} = \frac{GS_{11}^{nn}}{2} = \frac{0.5}{2} = 0.25$$

- Progetto 2

$$GS_{21} = \frac{GS_{21}^{nn}}{2} = \frac{1.35}{2} = 0.675$$

Il risultato della fase 2 risulta allora:

Gruppo G_1
Progetto 2, $GS_{21} = 0.675$ Progetto 1, $GS_{11} = 0.25$

Fase 3: Identificazione e adattamento dei rischi

Per procedere all'identificazione dei rischi del nuovo progetto si effettua uno scorrimento della tabella dei progetti estratti gruppo per gruppo, esaminando i rischi presenti nei progetti che ne fanno parte. Ad ogni rischio incontrato si associa un numero indicativo dell'appartenenza di tale rischio al nuovo progetto, detto *grado di appartenenza* GA_h (dove il rischio $Risk_h$ è un rischio di specifiche note al sistema incontrato durante lo scorrimento).

GA_h viene usato come contatore per valutare quanto è giusto che il rischio $Risk_h$ appartenga al nuovo progetto, esso viene sempre incrementato a partire da 0 per ogni rischio. Maggiore è il valore finale di GA_h e più l'esperienza passata ci suggerisce che il rischio si ripresenta anche nel nuovo progetto.

Ogni volta che si incontra il rischio $Risk_h$ nello scorrimento della struttura tabellare si procede a incrementare GA_h come segue:

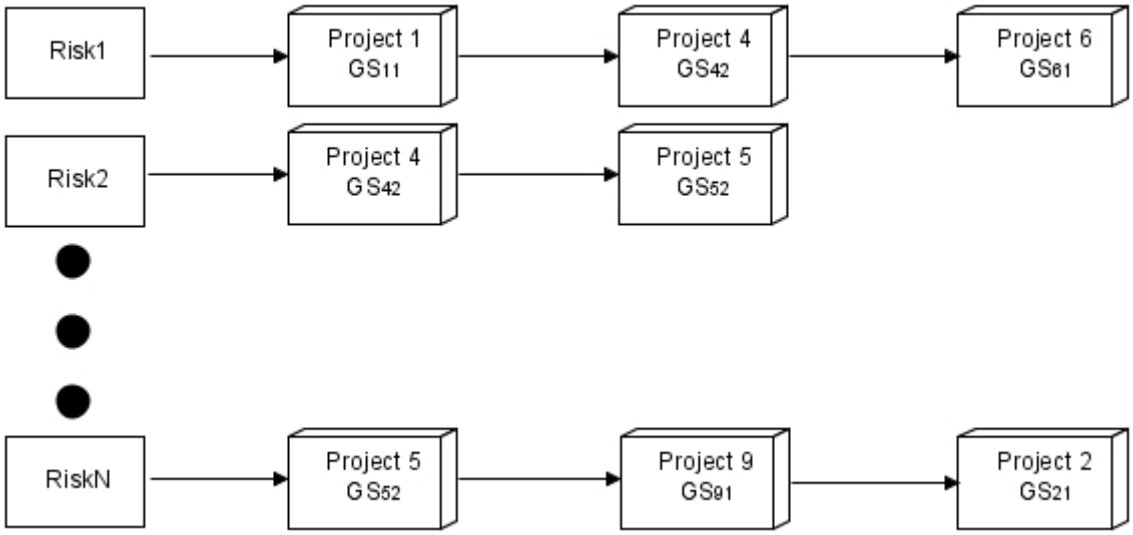
$$GA_h = GA_h + GS_{ij}$$

dove appunto si va a sommare al grado di appartenenza del rischio il grado di somiglianza del progetto incontrato (nell'ottica del gruppo che si sta esaminando in questo momento).

In questo modo, al termine dello scorrimento, il GA_h ci da un rank del rischio che tiene conto sia del *grado di somiglianza dei progetti in cui il rischio si è presentato* sia della *ricorrenza di tale rischio all'interno della tabella dei progetti estratti*.

Durante lo scorrimento della tabella dei progetti estratti, per ogni rischio incontrato, si memorizzano in una opportuna struttura in memoria i progetti in cui tale rischio è stato incontrato nella tabella dei progetti estratti, in questo modo le operazioni che seguono risultano più semplici ed è facile per ogni rischio recuperare i progetti che è necessario esaminare. In particolare, per ogni rischio si memorizzano gli elementi della tabella dei progetti estratti in cui è stato incontrato (e dunque sia il progetto in cui compare sia il grado di somiglianza del progetto stesso nell'ambito di un particolare gruppo).

La struttura che si viene a creare prende il nome di *struttura rischio - progetto* e la sua forma può essere ad esempio la seguente:



Notiamo che in questa struttura un progetto può presentarsi più volte associato ad un unico rischio (oltre che ovviamente associato a diversi rischi): questo effetto deriva dal fatto che lo stesso progetto (avente un certo rischio) può venire incontrato più volte nella struttura tabellare come appartenente a diversi gruppi. La situazione appena presentata non è assolutamente dannosa, essa può invece essere utilizzata a nostro vantaggio in quanto la presenza molteplice di un solo rischio deriva da una somiglianza al nuovo progetto sotto diversi aspetti (gruppi) e dunque questo può portare le somme pesate che seguono verso il modello adottato per il progetto ridondante nella struttura.

Passiamo ad esaminare come viene fatto l'adattamento dei rischi recuperati, notiamo che comunque non si è ancora fatta una selezione dei rischi da aggiungere effettivamente al nuovo progetto.

Per ognuno dei rischi, si scorrono i progetti ad esso associati usando la struttura rischio - progetto.

Lo scorrimento viene fatto per ogni rischio per adattare i parametri dei casi passati al caso in esame, per fare questa operazione in pratica si fa una media pesata sia per la probabilità *iniziale* del rischio nel nuovo progetto sia per il suo impatto *iniziale*.

Un progetto può presentarsi più volte per un unico rischio (come già spiegato) in quanto esso viene aggiunto alla struttura sopra semplicemente quando nella lettura di tale progetto nell'ottica di un gruppo si trova il rischio, dunque se un progetto appartiene a più gruppi allora sicuramente comparirà più volte nella struttura sopra associato ad un unico rischio.

Indichiamo adesso con $Proj_f^h$ l'f-esimo elemento che compare nella lista interna alla struttura rischio - progetto associata al rischio $Risk_h$. Definiamo $Prob_f^h$ la probabilità iniziale del rischio che compare nel progetto $Proj_f^h$ per il rischio $Risk_h$, e $Impa_f^h$ il suo impatto. Il progetto inserito nella lista ($Proj_f^h$) è un elemento che fa anche parte della tabella ricavata al termine della fase 2, per cui è disponibile il grado di somiglianza GS_{ij} (dove G_j è il gruppo da cui è stato estratto), rinominiamo adesso questo termine $GS_f = GS_{ij}$.
Chiamiamo infine $Prob_h$ la probabilità iniziale adattata per il rischio $Risk_h$ nel nuovo progetto, e $Impa_h$ l'impatto adattato.
In definitiva:

$$\begin{cases} Prob_h = \dfrac{\sum_f (Prob_f^h * GS_f)}{\sum_f GS_f} \\ Impa_h = \dfrac{\sum_f (Impa_f^h * GS_f)}{\sum_f GS_f} \end{cases}$$

Dunque si fa una media pesata dove in ogni si da un peso ad ogni probabilità e impatto dato dalla somiglianza del progetto di cui il rischio fa parte; inoltre, se un rischio appartiene ad un progetto somigliante sotto diversi aspetti al nuovo progetto (in diversi gruppi), si da più peso ai valori del rischio in tale progetto in quanto il termine relativo a tale progetto compare più volte nella sommatoria anche se eventualmente con diverso grado di somiglianza (relativo a diversi gruppi).

A questo punto abbiamo dei rischi che possono o meno appartenere al nuovo progetto, ognuno con un valore di impatto e probabilità adattati al nuovo progetto e un grado di appartenenza al nuovo progetto GA_h.

Esiste l'eventualità che tutti i rischi con grado di appartenenza molto alto siano in realtà poco rilevanti in quanto il loro impatto e la loro probabilità sono veramente bassi.
Per evitare di considerare nel nuovo progetto rischi di questo tipo, si calcola per ogni rischio un <u>indice di rischio</u> RI (risk index) che misura la rilevanza del rischio all'interno del nuovo progetto.
Sia allora RI_h l'indice di rischio del rischio h-esimo come definito precedentemente:

$$RI_h = Impa_h[in\ euro] * \frac{Prob_h}{100}$$

dove $Impa_h$ e $Prob_h$ sono impatto in euro e probabilità iniziali del rischio adattati al nuovo progetto, come appena calcolati.
Si scartano a questo punto tutti i rischi aventi indice di rischio inferiore al valore MIN_RI_THRESHOLD (configurabile dall'esperto nel file di configurazione init.conf e tramite interfaccia grafica).

Per evitare che al nuovo progetto siano comunque associati troppi rischi e per alleggerire il processo di inferenza, si selezionano a questo punto solamente i primi al più MAX_RISKS_PER_PROJ rischi con grado di appartenenza più elevato (dove MAX_RISKS_PER_PROJ è un parametro di inferenza configurato dall'utente nel file init.conf e tramite GUI).

MAX_RISKS_PER_PROJ rischi possono essere ancora troppi rischi per l'azienda da gestire in determinati progetti, dunque per lasciare all'utente un pò di voce in capitolo viene presentata tramite interfaccia grafica una barra di scorrimento normalizzata al 100% che permette all'utente stesso di spostare l'indicatore per scegliere un diverso numero di rischi da aggiungere al nuovo progetto in base al grado di appartenenza che si vuole considerare (es. 100% tutti i rischi, 0% solo il rischio a grado di appartenenza massimo).
Le descrizioni dei rischi aggiunti al nuovo progetto, le cause e gli effetti verranno copiati dagli omonimi campi nelle specifiche dei rischi disponibili al sistema, una specializzazione dei valori può avvenire solo per mano dell'utente.

Per semplificare le fasi successive, si eliminano dalla struttura rischio - progetto *tutti i rischi che non sono stati inseriti nel nuovo progetto*. In essa rimangono solo i rischi effettivamente presenti nel nuovo progetto legati agli elementi nella tabella dei progetti estratti in cui tali rischi sono stati individuati.

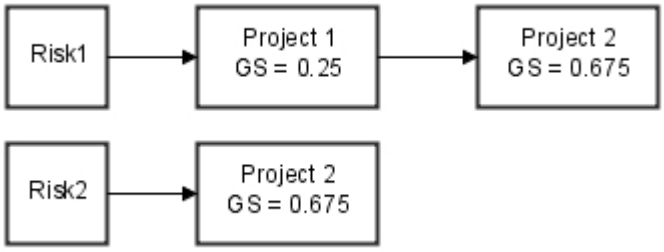
Come si capisce da quanto sopra, *non è possibile che nel nuovo progetto vengano inserite molteplici istanze dello stesso rischio* in quanto nello scorrimento della tabella dei progetti estratti ogni volta che si incontra un rischio già incontrato esso viene riconosciuto come rischio già esaminato: si aggiunge dunque l'elemento alla struttura rischio - progetto nella lista relativa al particolare rischio, modificando il GA_h già inizializzato e non creandone uno nuovo.

ESEMPIO

Osservando l'esempio di CBR in esame abbiamo:

- Rischio 1
  $GA_1 = GS_{11} + GS_{21} = 0.925$
- Rischio 2
  $GA_2 = GS_{21} = 0.675$

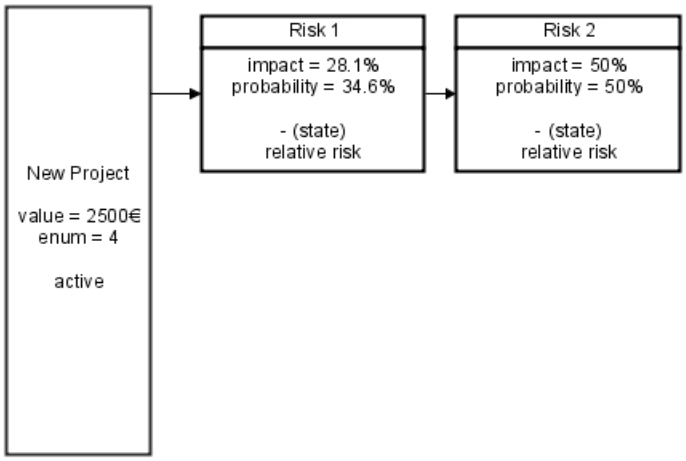
La struttura rischio - progetto assume la seguente forma:



L'adattamento di impatto e probabilità nei due casi produce:

- Rischio 1
  $Prob_1 \approx 34.6\%$ (ossia una probabilità spostata verso quella nel progetto 2)
  $Impa_1 \approx 28.1\%$ (ossia un impatto spostato verso quello nel progetto 2)
  $RI_1 \approx value * \frac{Impa_1}{100} * \frac{Prob_1}{100} \approx 243.06€$
- Rischio 2
  $Prob_1 = 50\%$ (pari a quella nel secondo progetto, è il nostro unico dato)
  $Impa_1 = 50\%$ (pari a quello nel secondo progetto, è il nostro unico dato)
  $RI_2 \approx value * \frac{Impa_2}{100} * \frac{Prob_2}{100} = 625.00€$

Ipotizziamo che nessuno dei due rischi venga filtrato grazie al controllo fatto su MIN_RI_THRESHOLD.
Ipotizziamo infine che entrambi i rischi vengano selezionati dall'utente come rischi presenti nel nuovo progetto.
Il risultato finale della fase 3 risulta allora:



Fase 4: Identificazione e adattamento delle azioni di mitigazione e recovery

Si opera in modo diverso per recuperare e adattare le azioni di mitigazione associate al nuovo progetto e per recuperare e adattare le azioni di recovery associate al nuovo progetto.

Azioni di mitigazione

Sfruttiamo nuovamente la struttura (a questo punto ripulita dai rischi che non fanno parte del nuovo progetto) che collega un rischio identificato e adattato per il nuovo progetto ai progetti che lo contengono nella tabella dei progetti estratti.
Siano $\{Mit_Act_x\}$ le azioni di mitigazione note al sistema CBR di gestione del rischio, si associa ad ogni azione un valore W^x pari inizialmente a 0. Si scorre poi tutta la struttura rischio - progetto risultante dalla fase precedente e si opera come segue:

Per ogni rischio $Risk_h$

Per ogni elemento $Proj_f^h$ della lista associata al rischio $Risk_h$

Si scorrono tutte le azioni di mitigazione usate in $Proj_f^h$ per gestire il rischio $Risk_h$, e si procede come segue:

Caso 1.

Se:

- lo stato dell'azione *non* è chiuso

  oppure
- l'efficacia media (sulla probabilità e l'impatto del rischio $Risk_h$)[3] *non* è superiore ad una soglia MIN_MID_EFFECTIVENESS (parametro di inferenza configurato dall'utente nel file init.conf o da GUI, es 25%)

Si inserisce l'azione in una lista, modificandone il peso ogni volta che viene incontrata e cade nel presente caso 1 (compresa la prima volta).

Per le operazioni di modifica del peso si usa anche il grado di somiglianza del progetto $Proj_f^h$ in cui l'azione è stata incontrata, anch'esso si legge direttamente nella struttura rischio - progetto, chiamiamolo GS_f^h.

Si usa nel calcolo del peso anche quanto il rischio che l'azione di mitigazione va a combattere è effettivamente rilevante all'interno del nuovo progetto (per evitare che si dia troppo peso ad azioni che hanno forti efficacie ma su rischi non molto rilevanti nel nuovo progetto). Ricordiamo allora l'indice del rischio (priorità del rischio nel nuovo progetto)

$$RI_h = Impa_h[in\ euro] * \frac{Prob_h}{100}$$

Si procede dunque come segue:

- Se l'efficacia media è inferiore ad una soglia MAX_MID_INEFFECTIVENESS cambiata di segno (specificata dall'utente in init.conf o da GUI, es. 3%) si somma $(-10 * GS_f^h * RI_h)$ al peso di tale azione.

[3] In questo contesto la efficacia mediata è data semplicemente da:

$$\frac{impact\ effectiveness + probability\ effectiveness}{2}$$

dove ricordiamo che *impact_effectiveness* è sempre un valore percentuale.

- Se l'efficacia mediata è inferiore o pari allo 0% si somma ($-3 * GS_f^h * RI_h$) al peso di tale azione. Questo si fa perché fintanto che un'azione ha un trend negativo in alcuni casi, ma non pesantemente, è giusto penalizzarla meno di una (come nel caso precedente) che presenta una forte azione negativa.
- Se lo stato dell'azione è non chiuso ma l'efficacia mediata (tra probabilità e impatto) è superiore a MIN_MID_EFFECTIVENESS, si somma al peso di tale azione W^x il valore $3 * GS_f^h * RI_h$.
- Se lo stato dell'azione è chiuso ma l'efficacia mediata non è superiore a MIN_MID_EFFECTIVENESS si somma $2 * GS_f^h * RI_h$ al peso di tale azione W^x.
- Se lo stato dell'azione è non chiuso e l'efficacia mediata non è superiore a MIN_MID_EFFECTIVENESS (ma positiva) si somma $GS_f^h * RI_h$ al peso di tale azione.

Ricapitolando, possiamo fare una tabella dei casi per le azioni di mitigazione incontrate nel caso 1, dove:

- MME = MIN_MID_EFFECTIVENESS
- MMI = MAX_MID_INEFFECTIVENESS

<table>
<tr><th colspan="2" rowspan="2">$W^x = W^x +$</th><th colspan="4">Efficacia Media</th></tr>
<tr><th>≤ -MMI</th><th>> -MMI & ≤ 0</th><th>> 0 & ≤ MME</th><th>> MME</th></tr>
<tr><td rowspan="2">Stato</td><td>chiusa</td><td rowspan="2">$-10 * GS_f^h * RI_h$</td><td rowspan="2">$-3 * GS_f^h * RI_h$</td><td>$2 * GS_f^h * RI_h$</td><td>Caso 2
Caso 3</td></tr>
<tr><td>non chiusa</td><td>$GS_f^h * RI_h$</td><td>$3 * GS_f^h * RI_h$</td></tr>
</table>

Caso 2.

Se :

- lo stato dell'azione è chiuso
- l'efficacia media dell'azione è superiore a MIN_MID_EFFECTIVENESS
- lo stato del rischio $Risk_h$ nel progetto $Proj_f^h$ <u>non</u> è in uno dei due stati closed (closed_verified o closed_unverified)

Si seleziona immediatamente l'azione per l'aggiunta al nuovo progetto (se non presente[4]), che avverrà legandola al rischio $Risk_h$ in esame (se non già legata).

[4] Non abbiamo nessun problema ad evitare l'inserimento multiplo di una singola azione di mitigazione in quanto ogni volta che essa viene incontrata per mitigare un certo rischio è sufficiente fare un controllo sulle azioni già esaminate (memorizzate nella struttura azione di mitigazione - rischio, vedi oltre).

Caso 3.

Se:

- lo stato dell'azione è chiuso
- l'efficacia media dell'azione è superiore a MIN_MID_EFFECTIVENESS
- lo stato del rischio $Risk_h$ nel progetto $Proj_f^h$ è in uno dei due stati closed

Si procede come segue:

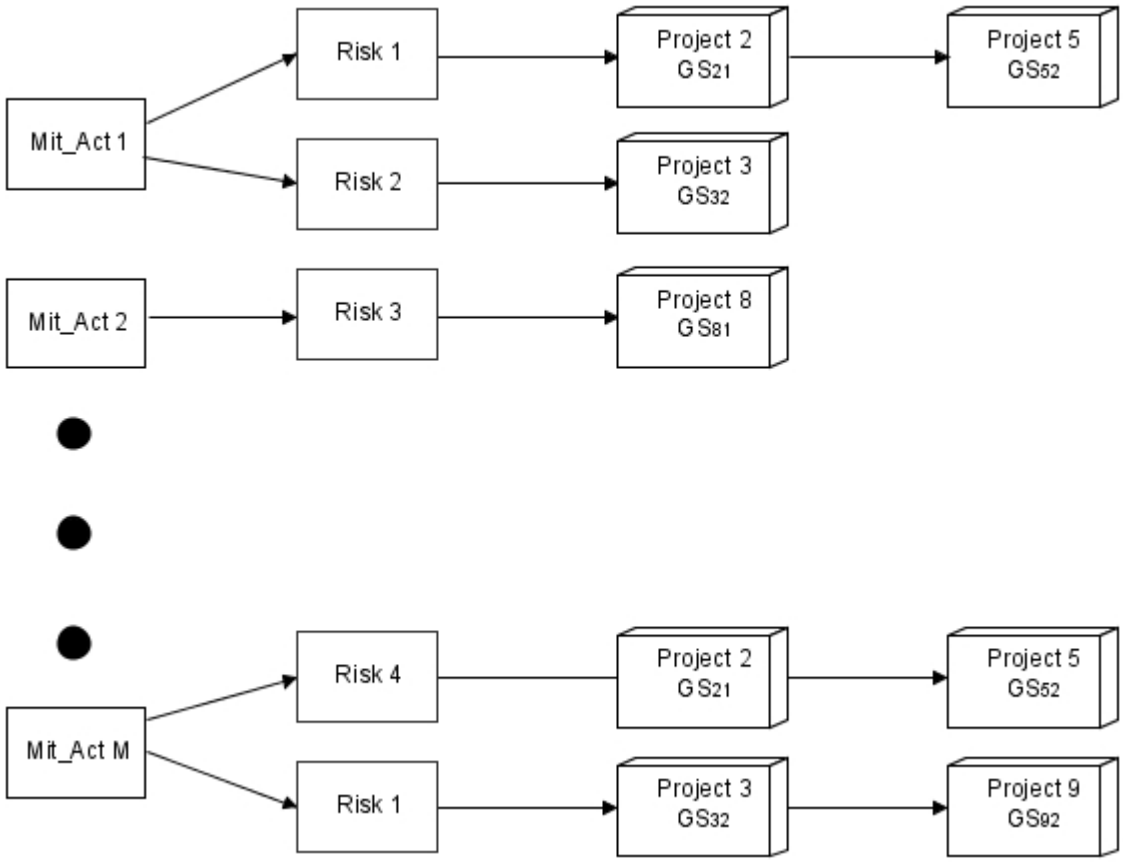
- Se l'azione ha avuto un effetto registrato nello storico del rischio $Risk_h$ a cui è collegata allora ha agito sul rischio prima che questo venisse chiuso e si opera come nel caso 2.
- Se l'effetto dell'azione non compare nello storico del rischio $Risk_h$ allora essa è stata completata e chiusa dall'utente solo dopo che il rischio è stato chiuso. Il motivo per cui risulta ancora collegata al rischio è che il sistema ritiene utile conservare informazioni su azioni che sono state previste dall'utente come mitiganti rischi che magari si sono chiusi prima che l'azione venisse svolta. In questo caso però si da meno peso all'azione considerando il suo stato non chiuso e si opera come descritto nel caso 1.

Alla fine dello scorrimento allora, avendo associato un peso a tutte le azioni con efficacia media sotto la soglia o stato non chiuso (oppure stato chiuso ma che non hanno avuto effetto su un rischio), si eliminano tutte le azioni che hanno peso risultante negativo o nullo, dopo di che si normalizzano i pesi al 100% (il che significa che il massimo peso rappresenta il 100%, e tutti gli altri vengono di conseguenza).
Dopo aver effettuato la normalizzazione, si presenta all'utente un'interfaccia grafica tramite cui è possibile scegliere il numero di azioni aggiuntive (rispetto a quelle già selezionate per l'inserimento in quanto aventi efficacia media sopra la soglia e chiuse con effetto su rischi) da considerare (in generale come valore di default si impone 15% come peso normalizzato minimo per accettare un'azione aggiuntiva). L'utente sceglie il peso normalizzato minimo per l'inserimento nel nuovo progetto in modo intuitivo, attraverso una barra di scorrimento.

Le azioni che risultano rientrare nei canoni richiesti vengono aggiunte al nuovo progetto legandole con tutti i rischi che esse sono in grado di gestire, questo avviene in modo semplice costruendosi (durante lo scorrimento della struttura rischio-progetto) una struttura che lega ogni azione di mitigazione incontrata a:

- rischi che è in grado di mitigare nel nuovo progetto
- *per ogni rischio che è in grado di mitigare* → elementi nella tabella dei progetti estratti dove l'azione viene usata per mitigare gli effetti del particolare rischio.

Questa struttura prende il nome di *struttura azione di mitigazione - rischio*, l'immagine sotto riportata ne è un esempio:



In pratica questa struttura si costruisce durante lo scorrimento della struttura rischio-progetto, aggiungendo tutte le azioni che vengono incontrate (anche se poi non verranno aggiunte al nuovo progetto). Al termine della scelta delle azioni si ripulisce però la struttura dalle azioni di mitigazione che non sono state scelte per il nuovo progetto.
Nella struttura sopra si trovano, dopo la ripulitura, sia le azioni che sono state scelte immediatamente per l'inserimento, sia le altre azioni che alla fine sono state aggiunte al nuovo progetto.

Ad ogni azione sono associati diversi rischi nella struttura, per ogni rischio sono disponibili gli elementi nella tabella dei progetti estratti in cui tale azione compare legata al rischio in esame. Questi contengono il grado di somiglianza del progetto oltre che l'efficacia dell'azione per la mitigazione del rischio nell'ambito del progetto sia in termini di probabilità che di impatto (individuabile all'interno del progetto stesso).

Il motivo per cui si sceglie di usare la struttura rischio - progetto per l'estrazione delle azioni di mitigazione al posto tabella dei progetti estratti è il seguente: se in un progetto è presente un'azione di mitigazione che gestisce *solamente* un rischio che però non è stato aggiunto al nuovo progetto, è inutile considerare tale azione per andare a vedere se è il caso di aggiungerla. Tale azione va ignorata e infatti non viene incontrata se si effettua lo scorrimento come descritto sopra.
È ovviamente possibile evitare questo problema anche utilizzando la tabella dei progetti estratti, tuttavia bisogna complicare il processo di scorrimento.

Al termine di questa procedura abbiamo un insieme di azioni di mitigazione che sono state aggiunte al nuovo progetto e legate nell'ambito del progetto a *tutti* i rischi che sono in grado di mitigare.
Le descrizioni delle azioni di mitigazione nel nuovo progetto verranno copiate dall'omonimo campo nelle specifiche delle azioni disponibili al sistema, una specializzazione delle descrizioni può avvenire solo per mano dell'utente.

Per ogni azione di mitigazione risultante nell'ambito del nuovo progetto devono però essere adattati i parametri caratteristici: efficacia sulla probabilità di ogni rischio mitigato, efficacia sull'impatto di ogni rischio mitigato e costo dell'azione di mitigazione. Per fare questa operazione si distinguono i due casi:

a) <u>*Calcolo dell'efficacia in termini di probabilità e impatto sul rischio*</u>

Questi parametri non sono associati ad una particolare azione di mitigazione all'interno del nuovo progetto, invece essi risultano associati ad una particolare *coppia* rischio - azione di mitigazione nell'ambito del progetto.
Possiamo pensare allora di utilizzare per l'adattamento la struttura azione di mitigazione - rischio presentata sopra.

Siano adesso $\{Mit_Act_x\}$ le azioni di mitigazione presenti nella struttura (che poi sono tutte le azioni di mitigazione facenti parte del nuovo progetto). Si definisce $\{Proj_l^x\}$ l'insieme degli elementi nella tabella dei progetti estratti associati all'azione x nella struttura (corrispondente ad un *sottoalbero* nella figura presentata precedentemente, avente come radice l'azione Mit_Act_x). Sia invece $\{Risk_h^x\}$ l'insieme dei rischi presenti nel sottoalbero dell'azione Mit_Act_x, che corrisponde all'insieme delle liste nel sottalbero (rami nel sottoalbero).
Con $\{Proj_l^{xh}\}$ si intende l'insieme degli elementi nella tabella dei progetti estratti associati al rischio h-esimo all'interno dell'insieme $\{Risk_h^x\}$ appartenente all'azione Mit_Act_x (questi elementi costituiscono un ramo del sottoalbero nella figura precedente).
Sia $prob_eff_l^{xh}$ l'efficacia in termini di probabilità dell'azione Mit_Act_x sul rischio $Risk_h$ nell'elemento $Proj_l^{xh}$ dell'insieme $\{Proj_l^{xh}\}$. Sia inoltre $impa_eff_l^{xh}$ l'efficacia in termini di impatto.
Chiamiamo infine GS_l in questo contesto il grado di somiglianza del progetto in $Proj_l^{xh}$.
Quello che si vuole calcolare è $impa_eff_{xh}$ e $prob_eff_{xh}$, ossia l'efficacia in termini di impatto e probabilità dell'azione Mit_Act_x (x-esima) nella struttura sul rischio $Risk_h$ (h-esimo), dunque i parametri dell'oggetto di tipo *Proj_Mit_Bind* tra l'azione x-esima e il rischio h-esimo.

Per trovare il valore di efficacia in termini di probabilità per l'azione di mitigazione Mit_Act_x sul rischio $Risk_h$ nel nuovo progetto, si scorrono tutti gli elementi $\{Proj_l^{xh}\}$ e si provvede a calcolare la seguente media pesata:

$$prob_eff_{xh} = \frac{\sum_l prob_eff_l^{xh} * GS_l}{\sum_l GS_l} \quad \begin{array}{c} efficacia\ in\ probabilità\ della \\ azione\ x-esima\ nei\ confronti\ del \\ rischio\ h-esimo\ nel\ nuovo\ progetto \end{array}$$

Si capisce che se un progetto in cui l'azione x-esima è stata usata per mitigare il rischio h-esimo compare in diversi gruppi allora la sua $prob_eff_{xh}$ compare diverse volte nella sommatoria a numeratore e a denominatore e dunque la soluzione finale si sposta un po' (giustamente) verso quella adottata nel progetto ridondante.

Allo stesso modo avremo:

$$impa_eff_{xh} = \frac{\sum_l impa_eff_l^{xh} * GS_l}{\sum_l GS_l} \quad \begin{array}{c} efficacia\ in\ impatto\ della \\ azione\ x-esima\ nei\ confronti\ del \\ rischio\ h-esimo\ nel\ nuovo\ progetto \end{array}$$

Questo è sempre un valore percentuale e dunque non da problemi in quanto non assume significato diverso in diversi progetti.

b) <u>*Calcolo del costo delle azioni di mitigazione*</u>

Abbiamo già a disposizione tutte le azioni di mitigazione registrate per il nuovo progetto. Per calcolare il *costo* di ogni azione di mitigazione si provvede a calcolare una media pesata del costo dell'azione di mitigazione in tutti i progetti recuperati e sopravvissuti all'interno della tabella dei progetti estratti (vista al termine della fase 2) che la prevedono, come peso si usa al solito il grado di somiglianza GS_{ij} di ogni progetto.

Ricordiamo che il costo di una azione di mitigazione può essere un costo relativo al valore del progetto oppure un costo assoluto (a seconda del valore di *rel_act* nelle specifiche dell'azione di mitigazione). Facendo questa operazione di media pesata non abbiamo problemi qualunque sia il valore di rel_act in quanto, indipendentemente dal progetto, l'azione ha sempre l'uno o l'altro comportamento.

Sia allora $cost_r^x$ l'r-esimo costo riscontrato per l'azione di mitigazione x-esima scorrendo la tabella dei progetti estratti, si definisce GS_r il grado di somiglianza associato al progetto in cui il costo $cost_r^x$ per l'azione Mit_Act_x (x-esima) è stato riscontrato. Chiamiamo inoltre $cost_x$ il valore che si vuole recuperare, ossia il costo per l'azione Mit_Act_x aggiunta al nuovo progetto.
Si calcola semplicemente:

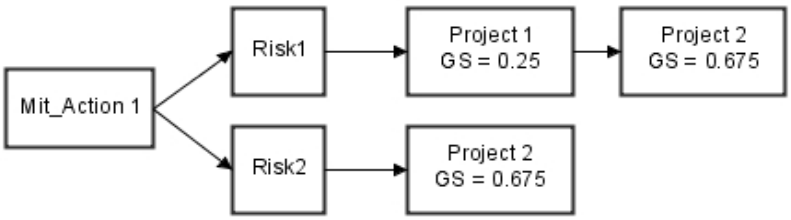
$$cost_x = \frac{\sum_r cost_r^x * GS_r}{\sum_r GS_r} \quad \begin{array}{c} costo\ della\ azione\ di\ mitigazione\ x-esima \\ nel\ nuovo\ progetto \end{array}$$

Al solito, è anche possibile che il costo di un'azione all'interno di un progetto estrattp si presenti più volte nella sommatoria a numeratore in quanto il progetto può comparire diverse volte all'interno della tabella dei progetti estratti se appartenente a gruppi diversi. Questo effetto collaterale è positivo e voluto in quanto se un progetto è stato ritenuto vicino al nuovo progetto in diversi aspetti è giusto che il suo peso sia preponderante sui progetti che invece non appartengono a questa categoria.

ESEMPIO

Vediamo che cosa succede nel semplice caso preso in esame fino ad ora.
Ipotizziamo che l'efficacia media riscontrata per l'unica azione presente nei due progetti risulti superiore alla soglia MME, per cui essa viene aggiunta direttamente al nuovo progetto.

La struttura azione di mitigazione - rischio risulta la seguente:

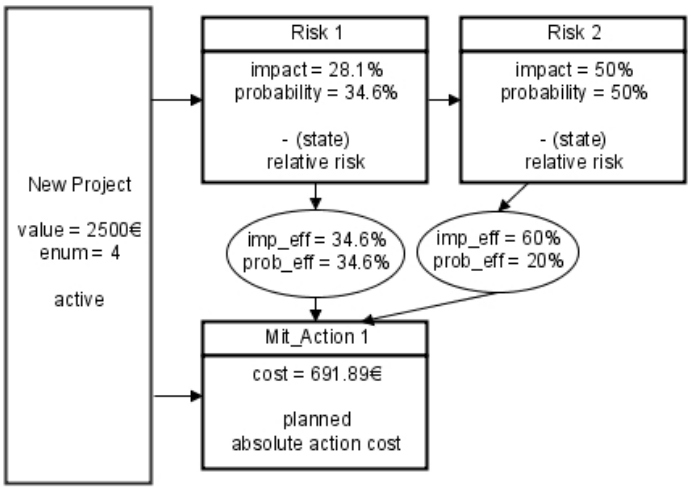


Dunque l'adattamento dell'efficacia dell'azione di mitigazione sui due rischi produce i seguenti risultati:

- Legame azione 1 rischio 1
  $prob_eff_{11} \approx 34.6\%$ (risultato spostato verso la situazione nel progetto 2)
  $impa_eff_{11} \approx 34.6\%$ (risultato spostato verso la situazione nel progetto 2)
- Legame azione 1 rischio 2
  $prob_eff_{12} = 20\%$ (quanto accade nel progetto 2, unico dato)
  $impa_eff_{12} = 60\%$ (quanto accade nel progetto 2, unico dato)

Osservando la tabella dei progetti estratti al termine della fase 2 si ricava inoltre:
$cost_1 \approx 691.89€$

E la situazione del nuovo progetto diventa la seguente.



Azioni di recovery

Sfruttiamo ancora la struttura rischio - progetto (ripulita dai rischi che non fanno parte del nuovo progetto) che collega un rischio individuato per il nuovo progetto ai progetti che lo contengono nella tabella dei progetti estratti.
Siano $\{Rec_Act_x\}$ le azioni di recovery note al sistema CBR di gestione del rischio, si associa ad ogni azione un insieme di pesi $\{W_h^x\}$ pari inizialmente a 0. In particolare ogni azione di recovery avrà un peso inizializzato a zero per ognuno dei rischi nella struttura rischio-progetto (pedice h).
Si scorre poi tutta la struttura rischio - progetto come segue:

Per ogni rischio $Risk_h$

Per ogni elemento della tabella dei progetti estratti $Proj_f^h$ associato al rischio $Risk_h$ (sia GS_f^h il grado di somiglianza con il nuovo progetto)

Si esamina (se presente) l'azione di recovery usata in $Proj_f^h$ per gestire il rischio $Risk_h$, e si procede come segue:

1. Se nell'ambito di $Proj_f^h$ l'azione Rec_Act_x compare per la gestione del rischio $Risk_h$ con risk_contingency inferiore a $VPart_f^h = original_RI_f^h * K$ (quello che si è chiamato *valore di partenza*[5]) allora $W_h^x = W_h^x + 2 * GS_f^h$ (azione buona).
2. Se nell'ambito di $Proj_f^h$ l'azione Rec_Act_x compare per la gestione del rischio $Risk_h$ con risk_contingency maggiore o uguale a $VPart_f^h = original_RI_f^h * K$ (valore di partenza) allora $W_h^x = W_h^x + GS_f^h$ (azione peggiore rispetto alla precedente).

In altri termini, ogni azione di recovery utilizzata in passato per un determinato rischio è candidata per diventare l'azione di recovery utilizzata nel nuovo progetto, però si da più peso a quelle usate in progetti simili al nostro e che hanno avuto una contingency necessaria più bassa (per gestire il rischio in esame) rispetto a quella attesa come media in azienda (stabilita al momento che si è deciso di gestire tale rischio con l'azione di recovery).

Ricapitolando, possiamo disegnare una semplice tabella:

	risk_contingency	
	< $VPart_f^h$	**≥** $VPart_f^h$
$W_h^x = W_h^x +$	$2 * GS_f^h$	GS_f^h

[5] Notiamo che $original_RI_f^h$ può essere un valore percentuale relativo al valore del progetto oppure un valore in euro. In entrambi i casi il calcolo resta quello sopra in quanto anche risk_contingency viene memorizzato allo stesso modo (ossia come un valore percentuale rispetto al valore del progetto se rel_risk è true).

Alla fine allora avremo associato diversi pesi ad ogni azione di recovery, uno per ogni possibile rischio presente nel nuovo progetto. Il peso è stato calcolato in modo tale da privilegiare le azioni che più spesso in passato sono riuscite a recuperare il rischio usando meno risorse di quelle medie usate per il recovery in azienda (risk_contingency < $VPart_f^h$) e in ogni caso considerando ogni valore come "pesato" col grado di somiglianza del progetto da cui è stato estratto.
Si considera dunque in questo modo:

- Ricorrenza dell'azione di recovery per un determinato rischio.
- Se l'azione di recovery è risultata buona (in grado di recuperare gli effetti del rischio spendendo meno della media).
- Quanto il progetto in cui si è utilizzata tale azione per il rischio in esame è simile al nuovo progetto.

Al termine dello scorrimento avremo associato ad ogni azione di recovery un peso per ognuno dei rischi presenti nel nuovo progetto.

Si scorrono a questo punto i rischi presenti nel nuovo progetto e per ogni rischio si aggiunge al nuovo progetto l'azione di recovery che risulta avere peso più elevato per tale rischio e *maggiore di 0* (un valore nullo indica che l'azione non è mai stata incontrata relativamente ad un certo rischio), legandola al rischio stesso.

Per il rischio h-esimo $Risk_h$ si aggiungerà l'azione:

$$Rec_Act_x : W_h^x = \max_x \{W_h^x\}$$

Se un'azione che dovrebbe essere aggiunta relativamente ad un rischio è già presente nel nuovo progetto allora significa che essa è stata aggiunta per recuperare un altro rischio, si provvede semplicemente in questo caso a creare una nuova connessione rischio - azione di recovery.

Può anche darsi che tutti i pesi relativi al rischio h-esimo siano nulli, questo significa che per il rischio h-esimo non sono state trovate azioni di recovery nella struttura rischio - progetto. In tal caso nel nuovo progetto al rischio h-esimo non viene associata alcuna azione di recovery (la strategia suggerita non comprenderà l'accettazione).

Ricordiamo comunque che, se l'associazione delle azioni di recovery ai rischi non fosse effettivamente quella voluta dall'utente, è sempre possibile cambiare i risultati in fase di validazione.

Le descrizioni delle azioni di recovery nel nuovo progetto verranno copiate dall'omonimo campo nelle specifiche delle azioni disponibili al sistema, una specializzazione delle descrizioni può avvenire solo per mano dell'utente.

Il campo original_RI presente nei legami instaurati tra rischi e azioni di recovery inserite nel nuovo progetto viene inizializzato banalmente col seguente valore (per l'azione di recovery x-esima legata al rischio h-esimo):

$$original_RI_{xh} = \frac{Prob_h}{100} * Impa_h$$

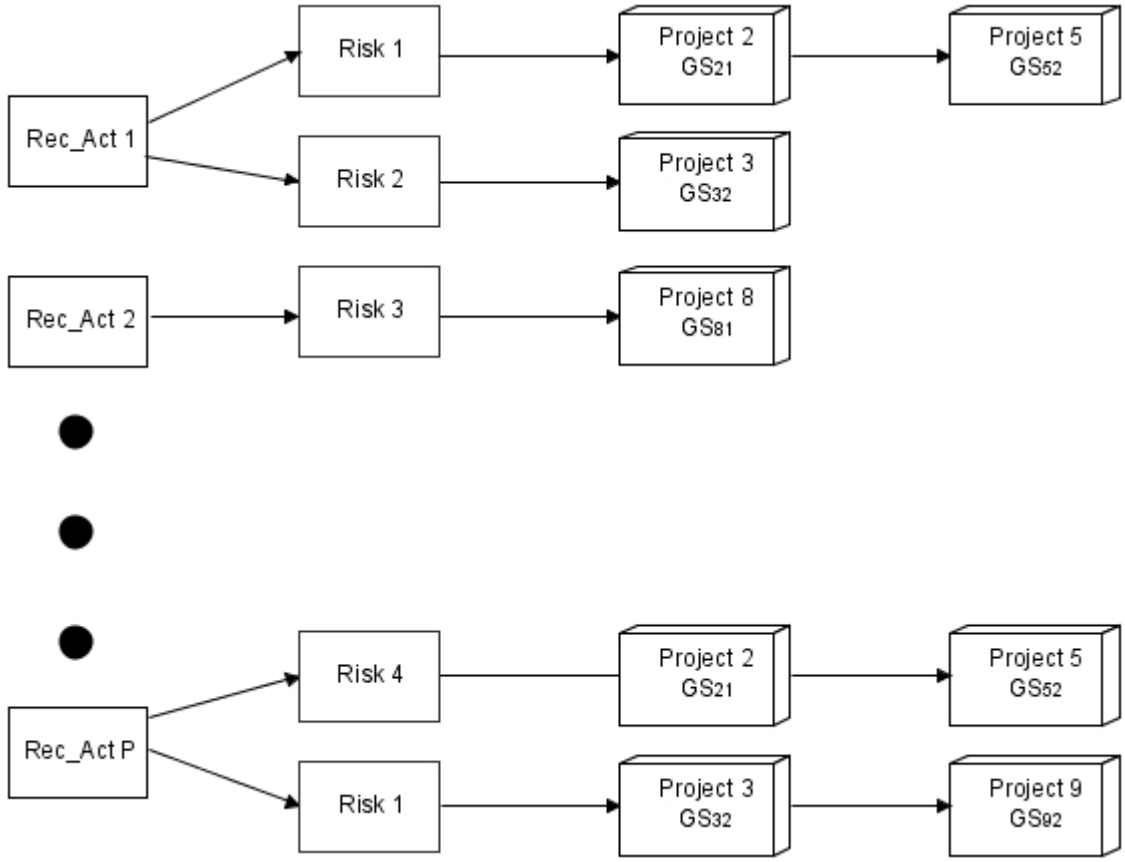
A questo punto rimane da vedere come funziona l'adattamento delle azioni di recovery che sono state recuperate e aggiunte al nuovo progetto. Notiamo che l'unico parametro che non conosciamo e in cui bisogna ancora scrivere un valore è il campo risk_contingency.

Il parametro risk_contingency non è associato ad una particolare azione di recovery all'interno del nuovo progetto, invece esso risulta associato ad una particolare *coppia* rischio - azione di recovery nell'ambito del progetto.

Ancora una volta conviene costruirsi una struttura durante la fase di scorrimento vista sopra, questa avrà una forma simile alla struttura azione di mitigazione - rischio.
È una struttura che lega ogni azione di recovery incontrata a:

- rischi che è in grado di recuperare nel nuovo progetto
- *per ogni rischio che è in grado di recuperare* → elementi nella tabella dei progetti estratti dove l'azione viene usata per recuperare gli effetti del particolare rischio.

Questa struttura prende il nome di *struttura azione di recovery - rischio*, l'immagine sotto riportata ne è un esempio:



Per questa valgono le stesse considerazioni che sono state fatte per la struttura azione di mitigazione - rischio.

La struttura sopra verrà ripulita, dopo lo scorrimento, da tutte le sottoliste corrispondenti a coppie rischio - azione di recovery che non sono state aggiunte al nuovo progetto come legami (Proj_Rec_Bind) e da tutti i sottoalberi corrispondenti a azioni di recovery che non sono state previste nel nuovo progetto (i criteri per l'inclusione sono stati visti sopra).

Siano adesso $\{Rec_Act_x\}$ le azioni di recovery presenti nella struttura azione di recovery - rischio (che poi sono tutte le azioni di recovery facenti parte del nuovo progetto). Si definisce $\{Proj_l^x\}$ l'insieme degli elementi nella tabella dei progetti estratti associati all'azione x nella struttura (corrispondente ad un *sottoalbero* nella figura presentata precedentemente, avente come radice l'azione Rec_Act_x). Sia invece $\{Risk_h^x\}$ l'insieme dei rischi presenti nel sottoalbero dell'azione Mit_Act_x, che corrisponde all'insieme delle liste nel sottalbero (rami nel sottoalbero).
Con $\{Proj_l^{xh}\}$ si intende l'insieme degli elementi nella tabella dei progetti estratti associati al rischio h-esimo all'interno dell'insieme $\{Risk_h^x\}$ appartenente all'azione Mit_Act_x (questi elementi costituiscono un ramo del sottoalbero nella figura precedente).
Sia $risk_contingency_l^{xh}$ la contingency relativa al rischio presente nell'elemento $Proj_l^{xh}$ dell'insieme $\{Proj_l^{xh}\}$ e $original_RI_l^{xh}$ il valore del campo original_RI presente nell'elemento $Proj_l^{xh}$.
Chiamiamo GS_l in questo contesto il grado di somiglianza del progetto in $Proj_l^{xh}$.

Si può allora ricavare per ogni elemento $Proj_l^{xh}$ al variare di *l* quello che è il valore atteso di risk_contingency valutando un comportamento medio aziendale rispetto al momento in cui l'azione di recovery è stata legata al rischio:

$$VPart_l^{xh} = original_RI_l^{xh} * K$$

Nel nuovo progetto è già noto il valore di:

$$original_RI_{xh} = \left(\frac{Prob_h}{100} * Impa_h\right)$$

$$VPart_{xh} = original_RI_{xh} * K$$

Dove il denominatore 100 è utilizzato per normalizzare le percentuali.

Dunque per stimare il valore di $risk_contingency_{xh}$ nel nuovo progetto si cerca di stimare la distanza che ci deve essere tra $risk_contingency_{xh}$ e $VPart_{xh}$, in questo modo si cerca di capire se l'azione tende ad avere un comportamento in genere migliore o peggiore della media aziendale.
Notare che sia $risk_contingency_{xh}$ che $original_RI_{xh}$ hanno un significato diverso a seconda del tipo di rischio (campo rel_risk) e lo stesso vale per $risk_contingency_l^{xh}$ e $original_RI_l^{xh}$ estratti dagli elementi in $\{Proj_l^{xh}\}$.
Tuttavia, tutti questi valori fanno riferimento allo stesso rischio (l'h-esimo) e dunque hanno tutti un valore percentuale oppure tutti un valore assoluto a seconda del valore di rel_risk per il rischio in esame, pertanto quanto segue resta valido in entrambi i casi.

Per calcolare il valore di $risk_contingency_{xh}$ (ossia la contingency dell'azione di recovery x-esima per gestire il rischio h-esimo nel *nuovo* progetto) basta scorrere tutti gli elementi $\{Proj_l^{xh}\}$ e fare una media pesata dei valori che vi si trovano per $risk_contingency_l^{xh} - VPart_l^{xh}$, dopo di che si somma il risultato a $original_RI_{xh} * K$.

$$risk_contingency_{xh} = \frac{\sum_l (risk_contingency_l^{xh} - VPart_l^{xh}) * GS_l}{\sum_l GS_l} + VPart_{xh}$$

$$risk_contingency\ della$$
$$azione\ x - esima\ nei\ confronti\ del$$
$$rischio\ h - esimo\ nel\ nuovo\ progetto$$

Il valore risultate della risk_contingency per ogni azione di recovery relativa ad un certo rischio potrebbe dunque essere un valore calcolato considerando l'impatto del rischio in euro oppure un valore calcolato considerando l'impatto del rischio come percentuale rispetto al valore del progetto. In entrambi i casi il procedimento sopra è valido e il sistema si incarica di presentare i valori all'utente sempre in euro deducendo il caso in cui ci troviamo dal campo rel_risk nelle specifiche del rischio.

Al termine di questa fase è inoltre possibile fissare la strategia di gestione di ogni rischio e lo stato iniziale di ogni rischio.

ESEMPIO

Vediamo che cosa nella situazione presa in esame fino ad ora.
L'unica azione di recovery presente nella struttura rischio – progetto è la Rec_action 1, essa sarà l'unica a venire inserita nel nuovo progetto per recuperare il rischio 1.

La struttura azione di recovery - rischio finale risulta la seguente:

Ipotizzando adesso K = 1.0 si ottiene:

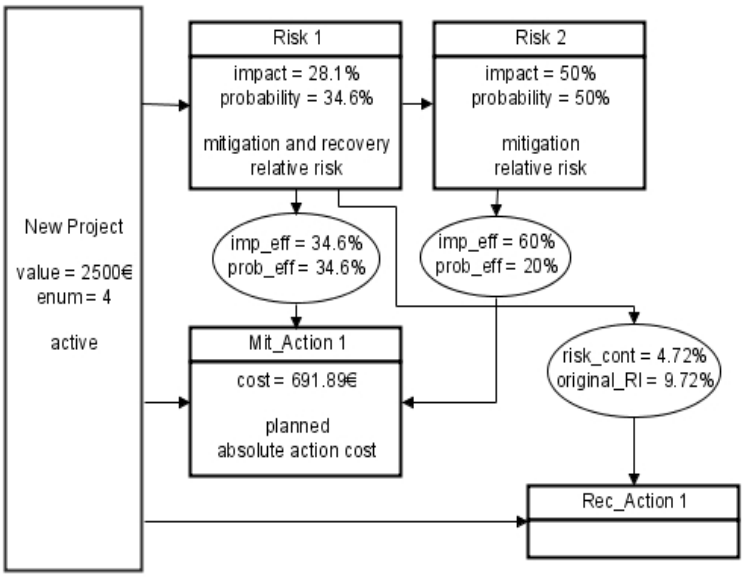
- Legame azione 1 rischio 1 (unico)

$$original_RI_{11} \approx 9.72\%$$
$$risk_cont_{11} \approx 4.72\%$$

Strategia di gestione e stato iniziale dei rischi:

- Rischio 1
  strategy = mitigation and acceptance
  stato = monitoring and recovery
- Rischio 2
  strategy = mitigation
  stato = monitoring

La situazione del nuovo progetto pre-validazione diventa la seguente.



E se l'utente non fa modifiche nella successiva fase di validazione, la struttura del nuovo progetto resterà tale al termine del Case Based Reasoning.

Fase 5: Validazione dei risultati del sistema

In questa fase si presenta il risultato dell'inferenza all'utente.
L'utente è in grado di modificare i risultati per adattarli a quello che la sua esperienza gli suggerisce.
Dopo la validazione il nuovo progetto viene memorizzato nella base di conoscenza come progetto attivo, e durante la vita del progetto il sistema provvederà ad assistere l'utente nella gestione dei rischi.
La memorizzazione permanente dei dati all'interno della base di conoscenza avviene però come già visto alla chiusura del progetto e dipende solamente da quanto sono pesanti le modifiche apportate dall'utente al risultato del sistema.
Il risultato della fase 4 viene allora *memorizzato nella base di dati fino alla chiusura del progetto*, separato dal progetto effettivo che è quello su cui lavora l'utente in fase di validazione e a posteriori.

Nella fase di validazione è possibile modificare liberamente il risultato del sistema.
Per presentare le funzionalità in modo agevole, si suddividono le possibilità di correzione dei risultati in diverse categorie:

- **categoria A**: eliminazione di componenti
- **categoria B**: modifica di componenti
- **categoria C**: aggiunta di componenti

categoria A

A1. *Eliminazione di interi rischi*

In questo caso si eliminano le azioni di mitigazione/recovery presenti nel progetto se utilizzate solo per uno dei rischi eliminati. Ovviamente si prevede anche a cancellare ogni legame che il rischio appena eliminato aveva con azioni di mitigazione o di recovery, cancellando anche i dati contenuti nelle strutture di legame.

A2. *Eliminazione di azioni di mitigazione*

È possibile in questo caso che uno o più rischi cambino strategia di gestione, in particolare:

- un rischio che era gestito solamente con tale azione di mitigazione (niente recovery) subisce una variazione della sua strategia di gestione in "spostamento".
- un rischio che era gestito invece con una combinazione di mitigazione e accettazione ma dove l'unica azione di mitigazione era quella che si è eliminata subisce una variazione della sua strategia di gestione in "accettazione".

Si eliminano ovviamente tutte le connessioni che collegavano l'azione appena eliminata con dei rischi. Può anche cambiare lo stato di uno o più rischi, come descritto in sezione 2.

A3. *Eliminazione di azioni di recovery*

In questo caso uno o più rischi cambiano strategia di gestione, in particolare:

- un rischio che era gestito solamente con tale azione di recovery (niente mitigazione) subisce una variazione della sua strategia di gestione in "spostamento".
- un rischio che era gestito con una combinazione di mitigazione e accettazione dove l'azione di recovery utilizzata è quella che si sta eliminando subisce una variazione della sua strategia in "mitigazione" (infatti ricordiamo che al più un'unica azione di recovery può essere utilizzata per gestire un rischio).

Si eliminano ovviamente tutte le connessioni che collegavano l'azione appena eliminata con dei rischi. Cambierà anche lo stato di uno o più rischi, come descritto in sezione 2.

A4. Eliminazione di legami rischio - azione di mitigazione

Per ogni azione di mitigazione prevista nel nuovo progetto, l'utente può modificare i rischi che tale azione è in grado di mitigare rimuovendo dei legami. L'eliminazione di un legame comporta che il rischio coinvolto può cambiare la sua strategia di gestione, in particolare:

- Se il rischio era gestito solamente con tale azione di mitigazione (niente recovery) subisce una variazione della sua strategia di gestione in "spostamento".
- Se il rischio era gestito invece con una combinazione di mitigazione e accettazione ma dove l'unica azione di mitigazione era quella che si è eliminata subisce una variazione della sua strategia di gestione in "accettazione".

Se l'azione coinvolta rimane a questo punto isolata da ogni rischio si provvede ad eliminare l'azione stessa come descritto sopra. Può anche cambiare lo stato del rischio, come descritto in sezione 2.

A5. Eliminazione di legami rischio - azione di recovery

Per ogni azione di recovery prevista nel nuovo progetto, l'utente può modificare i rischi che tale azione è in grado di recuperare rimuovendo dei legami. L'eliminazione di un legame comporta che il rischio coinvolto cambierà la sua strategia di gestione, in particolare:

- Se il rischio era gestito solamente con tale azione di recovery (niente mitigazione) subisce una variazione della sua strategia di gestione in "spostamento".
- Se il rischio era gestito invece con una combinazione di mitigazione e accettazione subisce una variazione della sua strategia in "mitigazione" (infatti ricordiamo che al più una unica azione di recovery può essere utilizzata per gestire un rischio con strategia di accettazione).

Se l'azione coinvolta rimane a questo punto isolata da ogni rischio si provvede ad eliminare l'azione stessa come descritto sopra. Cambierà anche lo stato del rischio, come descritto in sezione 2.

categoria B

B1 Variazione del valore di probabilità e impatto iniziali dei rischi del nuovo progetto

Ogni rischio può subire modifiche al valore di probabilità e/o al valore di impatto ricavato dal sistema, modifiche sul valore di impatto vengono effettuate in euro (indipendentemente dal valore di rel_risk) per maggiore intuitività dell'interfaccia.

B2 Variazione del valore di costo di un'azione di mitigazione

Ogni azione di mitigazione presente nel nuovo progetto può subire una variazione al suo valore di costo, modifiche da parte dell'utente vengono fatte sempre lavorando in euro indipendentemente dal valore di rel_act.

B3 Variazione del valore di efficacia in termini di probabilità e impatto di un'azione di mitigazione nei confronti di un rischio

L'utente può lavorare sui legami che collegano azioni di mitigazione e rischi, cambiando l'efficacia che un'azione ha su ogni rischio ad essa collegato (lavorando in questo caso con valori percentuali).

B4 *Variazione del valore di contingency di un'azione di recovery relativo ad un certo rischio*

All'utente è concesso di operare sui legami che collegano azioni di recovery e rischi, cambiando la contingency di una azione di recovery relativa ad ogni rischio ad essa legato, ossia il valore di risk_contingency (in questo caso l'utente lavora con valori monetari espressi in euro, il sistema effettua la traduzione nel caso in cui rel_risk sia true dividendo il numero inserito per il valore del progetto).

categoria C

C1 *Aggiunta di interi rischi*

L'utente può aggiungere rischi prelevandoli esclusivamente dalla lista dei rischi già noti, ossia scegliendo tra i rischi di cui sono già note le specifiche. È ovvio che all'utente è data la possibilità di aggiungere nuove specifiche di rischi da interfaccia grafica, come descritto nella sezione seguente.

Per ogni rischio aggiunto è possibile scegliere quali azioni tra le presenti nel progetto hanno un effetto di mitigazione e di recovery per tale rischio. A seconda delle scelte dell'utente si provvederà ad aggiornare opportunamente (come descritto in sezione 2) il campo strategia di gestione.

C2 *Aggiunta di azioni di mitigazione*

L'utente può aggiungere azioni di mitigazione prelevandole esclusivamente dalla lista delle azioni di mitigazione già note, ossia scegliendo tra le azioni di mitigazione di cui sono già note le specifiche. È ovvio che all'utente è data la possibilità di aggiungere nuove specifiche di azioni di mitigazione da interfaccia grafica, come descritto nella sezione seguente.

Per ogni azione aggiunta, l'utente deve decidere a quali rischi presenti nel progetto tale azione è associata in quanto ha su di essi un effetto di mitigazione. L'utente deve inoltre specificare il costo per l'azione di mitigazione nel progetto in esame e l'efficacia di tale azione su ogni rischio ad essa legato in termini di probabilità e impatto del rischio.

I rischi a cui l'azione è stata associata possono potenzialmente cambiare strategia di gestione, ad aggiornare tale campo ci pensa il sistema secondo la logica descritta in sezione 2.

C3 *Aggiunta di azioni di recovery*

L'utente può aggiungere azioni di recovery prelevandole esclusivamente dalla lista delle azioni di recovery già note, ossia scegliendo tra le azioni di recovery di cui sono già note le specifiche. È ovvio che all'utente è data la possibilità di aggiungere nuove specifiche di azioni di recovery da interfaccia grafica, come descritto nella sezione seguente.

Per ogni azione aggiunta, l'utente deve decidere a quali rischi presenti nel progetto tale azione è associata in quanto viene usata per recuperarne gli effetti. L'utente deve in questo caso specificare quale è la frazione di contingency dell'azione di recovery destinata a ciascun rischio stabilendo i valori dei campi risk_contingency presenti in ogni connessione. I valori dei campi original_RI non vengono specificati dall'utente in quanto calcolati dal sistema come specificato precedentemente e dunque utilizzando i valori attuali di impatto e probabilità dei rischi.

I rischi a cui l'azione è stata associata possono potenzialmente cambiare strategia di gestione, ad aggiornare tale campo ci pensa il sistema secondo la logica descritta in sezione 2.

C4 *Aggiunta di legami rischio - azione di mitigazione*

Per ogni azione di mitigazione prevista nel nuovo progetto, l'utente può modificare i rischi che tale azione è in grado di mitigare, aggiungendo dei legami.

Nel caso in cui l'utente voglia creare dei nuovi legami, egli è obbligato a specificare l'efficacia dell'azione in questione in termini di probabilità e impatto sul rischio al quale questa azione si sta legando (come valori percentuali).

C5 *Aggiunta di legami rischio - azione di recovery*

Per ogni azione di recovery prevista nel nuovo progetto, l'utente può modificare i rischi in cui tale azione viene usata per recuperare gli effetti, aggiungendo dei legami.

Nel caso in cui l'utente comandi una simile operazione, è però obbligato anche a definire il nuovo valore per la contingency di quella azione destinata a recuperare il rischio a cui è stata legata (valore di risk_contingency) lavorando in euro. Il valore del campo original_RI non viene specificato dall'utente in quanto calcolato dal sistema come specificato precedentemente e dunque utilizzando i valori attuali di impatto e probabilità del rischio.

Alla fine della validazione la **revision (revisione)** del progetto è lasciata pari ad 0 indipendentemente dalle modifiche effettuate (infatti il progetto non è ancora cominciato e non avrebbe senso aumentare la revisione fin da ora).

Questo conclude la trattazione delle 5 fasi del processo d'inferenza, la procedura appena mostrata è quella che viene seguita come conseguenza dell'inserimento di un nuovo progetto ad opera dell'utente.

Dobbiamo adesso concludere la specifica del ruolo del modulo di inferenza con un ultimo aspetto fondamentale.

4.2 Chiusura di progetti e aggiunta permanente di casi

Un progetto inserito dall'utente tramite interfaccia grafica viene sempre memorizzato nella base di conoscenza dopo aver subito inizialmente il processo di inferenza da cui si ricavano rischi e azioni suggerite per il risk management di tale progetto e la validazione. Dunque un progetto in corso è sempre presente nella base di conoscenza come caso da usare per le inferenze che avvengono mentre il progetto è attivo.

Quando però un progetto viene chiuso si ha che si valuta se tenerlo nella knowledge base cercando di capire se l'utente ha fatto modifiche significative al risultato del sistema aggiungendo conoscenza utile per le inferenze successive. Questo compito è affidato al modulo di inferenza.

Ogni volta che l'utente inserisce un nuovo progetto, il sistema lancia il processo di inferenza che restituisce rischi e azioni per il nuovo progetto con tutti i parametri del caso. L'utente può poi modificare i risultati tramite una prima fase detta di validazione al termine dell'inferenza (vista precedentemente), oppure post-validazione (come descriveremo successivamente).

Il modulo di inferenza, attraverso il modulo di gestione della conoscenza, *salva una copia dell'intero progetto così come il sistema lo avrebbe gestito in modo automatico*, ossia si salvano i risultati del processo di inferenza prima della fase di validazione.
Si tiene inoltre la copia separata dall'effettivo progetto su cui si opera l'utente.
Questo viene fatto creando in pratica non uno, ma ***due*** progetti.

Il Project_ID del progetto copia, progetto che viene conservato dal sistema fino alla chiusura del progetto di cui è la copia iniziale, viene identificato tramite una tabella apposita presente nella base di conoscenza che è la tabella *proj_copies* (vista in sezione 3).

Quando l'utente comanda la chiusura di un progetto aperto, il sistema carica il project_ID del progetto copia leggendo nel campo copy_ID della tabella di cui sopra, ed inizializza a zero un contatore che chiamiamo ***mod_count***. Dopo di che comincia ad esaminare il progetto copia e il progetto in chiusura, ed opera come segue[6]:

- Se un rischio presente nel progetto copia non è presente nel progetto in chiusura allora:
  $mod_count = mod_count + 100$
- Se un rischio presente nel progetto in chiusura non è presente nel progetto copia allora:
  $mod_count = mod_count + 100$
- Se un'azione presente nel progetto copia (sia essa di mitigazione o di recovery) non è presente nel progetto in chiusura allora:
  $mod_count = mod_count + 70$
- Se un'azione presente nel progetto in chiusura (sia essa di mitigazione o di recovery) non è presente nel progetto copia allora:
  $mod_count = mod_count + 70$
- Se un legame tra un'azione comune ai progetti (sia essa di recovery o di mitigazione) e un rischio comune ai progetti è presente nel progetto copia ma non nel progetto in chiusura allora:
  $mod_count = mod_count + 40$
- Se un legame tra un'azione comune ai progetti (sia essa di recovery o di mitigazione) e un rischio comune ai progetti è presente nel progetto in chiusura ma non nel progetto copia allora:
  $mod_count = mod_count + 40$
- Se in un rischio presente in entrambi i progetti i valori di probabilità *iniziale* differiscono allora, definendo con $probability^c$ (copia) la probabilità presente nel progetto copia e con $probability^e$ (effettivo) la probabilità nel progetto effettivo in chiusura, si ha che:
  $mod_count = mod_count + |probability^c - probability^e|$

[6] Tenere presente che, come già detto, *nella base di dati e in memoria i valori che in questa documentazione sono descritti come "percentuali" sono rappresentati come numeri reali da 0 a 100*, ad esempio:
- 100% -> 100.00
- 2% -> 2.00
- 34.56% -> 34.56

- Se in un rischio presente in entrambi i progetti abbiamo che i valori di impatto *iniziale* differiscono allora, definendo con $impact^c$ l'impatto presente nel progetto copia e con $impact^e$ l'impatto nel progetto effettivo in chiusura, si ha che:

$$mod_count = mod_count + \begin{cases} \frac{|impact^c - impact^e|}{impact^c} * 100 & if \quad rel_risk = false \ \& \ impact^c \neq 0 \\ 40 \quad if \quad rel_risk = false \ \& \ impact^c = 0 & (impact^e \neq 0) \\ |impact^c - impact^e| \quad if \quad rel_risk = true \end{cases}$$

  Infatti se rel_risk è vero allora il rischio ha un impatto relativo al valore del progetto e dunque gli impatti sono memorizzati come valori percentuali su cui si può fare una semplice differenza in modulo, se invece rel_risk è falso allora il rischio ha un impatto memorizzato in euro e si valuta la differenza come distanza riferita al valore iniziale (se l'impatto iniziale era nullo allora si aggiunge **40** a mod_count).
- Se in un'azione di mitigazione presente in entrambi i progetti abbiamo che i valori di costo differiscono allora, definendo con $cost^c$ il costo presente nel progetto copia e con $cost^e$ il costo nel progetto effettivo in chiusura, si ha che:

$$mod_count = mod_count + \begin{cases} \frac{|cost^c - cost^e|}{2 * cost^c} * 100 & if \quad rel_act = false \ \& \ cost^c \neq 0 \\ 25 \quad if \quad rel_act = false \ \& \ cost^c = 0 & (cost^e \neq 0) \\ \frac{|cost^c - cost^e|}{2} \quad if \ rel_act = true \end{cases}$$

  Infatti se rel_act è vero allora l'azione ha un costo relativo al valore del progetto e dunque i costi sono memorizzati come valori percentuali su cui si può fare una semplice differenza in modulo, se invece rel_act è falso allora l'azione ha un costo memorizzato in euro e si valuta la differenza come distanza riferita al valore iniziale (se il costo iniziale è nullo allora si aggiunge **25** a mod_count).

  Il fattore 2 al denominatore serve per bilanciare l'importanza del costo di un'azione di mitigazione, meno significativo di altri parametri come quelli valutati precedentemente.
- Se un'azione di mitigazione viene usata in entrambi i progetti per mitigare gli effetti di uno stesso rischio, ma i valori di efficacia in termini di probabilità sul rischio differiscono, allora si procede come segue:

$$mod_count = mod_count + |prob_eff^c - prob_eff^e|$$

- Se un'azione di mitigazione viene usata in entrambi i progetti per mitigare gli effetti di uno stesso rischio, ma i valori di efficacia in termini di impatto sul rischio differiscono, allora si procede come segue:

$$mod_count = mod_count + |impa_eff^c - impa_eff^e|$$

  Si usa la stessa formula del caso (efficacia in termini di probabilità) sopra in quanto l'efficacia in termini di impatto è anch'essa sempre un valore percentuale, come spiegato precedentemente.

- Se un'azione di recovery viene usata in entrambi i progetti per recuperare gli effetti di uno stesso rischio ma i valori di risk_contingency risultano diversi, si opera come segue.
  Definendo $risk_contingency^c$ quella nel progetto copia, $risk_contingency^e$ quella nel progetto effettivo in chiusura e $original_RI^c$ il valore del campo $original_RI$ nel progetto copia, allora:
  $$mod_count = mod_count + \begin{cases} \dfrac{|risk_contingency^c - risk_contingency^e|}{original_RI^c * K} * 100 & if \quad original_RI^c \neq 0 \\ 30 \quad if \quad original_RI^c = 0 \end{cases}$$
  Ossia si calcola la distanza in valore assoluto (che ovviamente ha un significato diverso a seconda del valore di rel_risk) e si riferisce il suo valore al valore di partenza per la contingency valutato nel progetto copia. Questa tecnica funziona qualunque sia il valore di rel_risk. Se il valore di partenza nel progetto copia è nullo allora si aggiunge a mod_count il valore **30**.

Notare che all'utente non è consentito di operare sul valore di original_RI presente sulle connessioni tra azioni di recovery e rischi, dunque i valori nei due progetti non potranno differire.

A questo punto abbiamo terminato di incrementare il campo mod_count e si procede come segue: se mod_count ≥ MOD_THRESHOLD (parametro di inferenza configurato dall'utente tramite init.conf o GUI, *valore consigliato 100*) il progetto stesso va mantenuto in memoria come caso utile all'inferenza in quanto si ritiene che abbia conoscenza nuova e significativa per migliorare il funzionamento del sistema (consentendogli di *apprendere*), altrimenti il progetto non è utile per le successive inferenze e viene flaggato nella base di dati come informazione da mantenere per la politica aziendale ma inutile ai fini del CBR (ossia il progetto non fa più parte della base di conoscenza).

Al termine del controllo incrociato sul progetto in chiusura e sulla sua copia ci sono allora due possibilità:

a. *mod_count ≥ MOD_THRESHOLD*
   Si elimina dalla base di dati il progetto copia e si lasciano le tuple relative al progetto appena chiuso nella base di conoscenza in quanto ritenute significative (costituiscono un caso utile).
b. *mod_count < MOD_THRESHOLD*
   Si rimuove dalla base di dati il progetto copia e si segnala che il progetto appena chiuso non è utilizzabile per il processo di inferenza (non fa più parte della base di conoscenza). Questo si fa perché la conoscenza in tale progetto non è ritenuta significativa in quanto inferibile (più o meno) dalla conoscenza già in possesso del sistema CBR).

Il motivo per cui si è deciso di tenere in memoria un progetto copia per tutti i progetti attualmente in corso è che questa tecnica consente di trattare facilmente sia modifiche in fase di validazione al termine dell'inferenza, sia modifiche fatte a posteriori (come l'aggiunta di un rischio). Una via

alternativa sarebbe contare solamente le modifiche quando le stesse vengono eseguite memorizzando nella base di conoscenza un contatore apposito, il problema però in questo caso è che sarebbe impossibile accorgersi di quando l'utente, fatta una modifica, torna indietro riavvicinandosi alla soluzione proposta dal sistema.

Vogliamo adesso fare un confronto tra il metodo di apprendimento utilizzato nel CBR Risk Management System appena esaminato e quello tipico che si trova nella letteratura riguardante la progettazione di Sistemi Esperti (SE) utilizzanti il ragionamento basato sui casi.
In generale l'apprendimento può avvenire in due modi:

1. Inserimento di nuovi casi nella base dei casi o base di conoscenza
2. Modifica delle politiche di estrazione e/o adattamento dei casi passati in funzione del successo o insuccesso di una soluzione proposta dal sistema CBR.

Il tipo di apprendimento che si è utilizzato per il CBR Risk Management System ricade ovviamente nella prima categoria.
Quando si usa questo tipo di apprendimento, la raccomandazione generale è quella di memorizzare nuovi casi quando questi *non sono troppo simili ad altri casi già presenti*. In questo modo si cerca di piazzare dei casi in modo approssimativamente uniforme all'interno dello spazio dei casi, al fine di ridurre la distanza media che si ha tra i casi recuperati e quello nuovo fornito dall'utente per migliorare la risposta del sistema.

La soluzione che si adotta nel CBR Risk Management System invece non memorizza un caso anche quando questo risulta distante dagli altri casi nella base di conoscenza, a meno che l'utente non faccia modifiche sostanziali al progetto correggendo il risultato del sistema.
Questa soluzione presenta alcuni vantaggi sostanziali che la rendono preferibile nel sistema in esame:

- Semplifica significativamente la <u>manutenzione del sistema</u>.
  Riempiendo in modo più o meno uniforme lo spazio dei possibili progetti, durante l'utilizzo del sistema potrebbero peggiorare le prestazioni grazie al grande numero di progetti che è stato memorizzato. Nella soluzione adottata dal CBR Risk Management System abbiamo che un caso viene memorizzato solamente in caso di errore significativo commesso dal sistema, per cui le prestazioni peggiorano solamente in caso di errore. Quanti più errori il sistema commette e tanti più casi sarà costretto a memorizzare per affinare l'inferenza, quando l'errore si sarà ridotto sotto la tolleranza dell'utente allora il sistema smetterà di apprendere nuovi casi (non peggiorando le prestazioni nel tempo).
  Per cui la soluzione proposta risulta un *compromesso tra capacità di apprendimento e necessità di manutenzione* del sistema.

- Gestisce meglio l'evoluzione delle politiche aziendali.
  Se nel tempo il comportamento dell'azienda subisce cambiamenti nelle preferenze di gestione dei rischi, un progetto simile ad un'altro può essere gestito in modo leggermente diverso se affrontato dopo molto tempo. Questo caso andrebbe memorizzato in quanto significativo per le future inferenze, ma basandoci solo sulla distanza da altri progetti esso verrebbe scartato dalla base di conoscenza.
  La soluzione adottata nel CBR Risk Management System invece consente di memorizzare il progetto in quanto (se la soluzione non era esattamente quella voluta dall'azienda) l'utente avrà modificato la risposta del sistema, forzando l'apprendimento.
  Lo stesso ragionamento si può fare qualora la situazione aziendale è tale per cui in casi tra loro molto simili si possono avere delle differenze nelle soluzioni dovuti a parametri sconosciuti al sistema CBR (*eccezioni*).

5 Modulo di Interfaccia

Scopo di questo modulo è occuparsi di tutti gli aspetti di interazione con l'utente, esso deve essere in grado di dialogare sia con il modulo di gestione della conoscenza che con il modulo di inferenza per svolgere correttamente il suo compito soddisfacendo le richieste che gli vengono inoltrate tramite interfaccia grafica.

Le interazioni con l'utente sono molteplici, in particolare si cerca nella seguente lista di riassumere tutte le possibili forme di interazione. Notiamo che l'utente ha (a posteriori della fase di validazione) la capacità di continuare a lavorare sui risultati dell'inferenza cambiandoli in massima parte in base alla sua esperienza passata e a quello che sta succedendo in azienda mentre il progetto viene portato avanti. Molte delle possibili modifiche di seguito elencate sono delle ripetizioni di quelle viste in fase di validazione: a rischio di risultare noiosi, cercheremo di essere il più precisi possibile.

Il sistema è provvisto di una elaborata interfaccia grafica che si occupa di semplificare l'interazione con l'utente e di renderla quanto più intuitiva possibile.

I punti da 1 a 6 descrivono utilità del sistema CBR per seguire l'utente nel processo aziendale di gestione dei rischi dei progetti.

1. *Aggiunta di (specifiche di) Azioni di mitigazione/recovery e di rischi*
   In questo caso all'utente è lasciata la specifica di ogni campo dell'oggetto che desidera aggiungere.

2. *Presentazione di un nuovo progetto al sistema*
   Un wizard intuitivo basato su interfacce grafiche guida l'utente attraverso le varie fasi dell'aggiunta di un progetto (le quali si ricollegano direttamente al processo di inferenza):
   - Specifica di tutte le features del nuovo progetto (ovviamente non si richiedono rischi e azioni che verranno ricavati dal sistema per inferenza)
   - Specifica degli attributi (features) da utilizzare per l'estrazione e ranking dei progetti simili e dei loro pesi nei due casi
   - Scelta del numero dei rischi di cui l'utente si vuole occupare
   - Scelta del numero di azioni di mitigazione aggiuntive di cui l'utente si vuole incaricare
   - Validazione

   Annullando l'operazione il progetto non viene aggiunto.

3. *Recupero di progetti in corso (e funzioni di ordinaria gestione)*

   Un progetto in corso può venir recuperato specificando uno o più tra i seguenti campi:

   - Project Name
   - Customer Name
   - Company ID

   Nelle query scritte per recuperare i progetti si provvede a fare una diversa selezione basata sugli attributi che l'utente ha specificato (i campi lasciati vuoti non verranno utilizzati). In ogni caso i progetti recuperati sono soltanto quelli con stato "active".

   Si presenta dunque all'utente un elenco di progetti che corrispondono a quanto richiesto, in questo elenco l'utente individua il progetto cercato.

   Quando si sceglie un progetto, vengono presentate tutte le informazioni disponibili su tale progetto, in questa interfaccia è inoltre presente una utilità per:

   - Chiudere le azioni di mitigazione dando eventualmente un feedback
     (la revisione del progetto in questo caso *non viene incrementata*)
   - Specificare che un'azione è cominciata spostando il suo stato su "active"
     (gli stati dei rischi cambiano automaticamente come necessario)
   - Chiudere dei rischi in quanto la loro gestione si è conclusa perché il loro effetto si è verificato (closed_verified) o per altri motivi (closed_unverified), come descritto in sezione 2.
   - Chiudere il progetto
     In questo caso si provvede a lanciare la procedura del modulo di inferenza che valuta se lasciare o meno il progetto nella base di conoscenza o rimuoverlo dalla stessa, come descritto nel paragrafo 4.2.
   - Modificare impatto e probabilità attuali di un rischio per cause non dipendenti dall'esecuzione di azioni di mitigazione, questa operazione causa l'aggiunta di un elemento nello storico del rischio. Si provvede inoltre all'*incremento della revisione del progetto*.

   È inoltre questa l'interfaccia che consente di operare sul progetto delle correzioni per la gestione opportuna dei rischi come descritto nei punti da 7 a 24.

4. *Browsing dei progetti chiusi*

   Vi è nel modulo di interfaccia un'utilità per il browsing di tutti i progetti chiusi (sia parte della base di conoscenza che fuori da essa). Si permette in questo modo di cancellare i singoli progetti qualora venissero considerati dall'esperto non più necessari. Si ricorda tuttavia che i progetti chiusi e memorizzati nella base di conoscenza costituiscono la base del processo di inferenza e dunque l'eliminazione di alcuni di essi è rischiosa in quanto può compromettere l'efficacia del Case Based Reasoning.

5. *Rebuild della base di dati*

   L'utente è in grado di comandare al sistema l'esecuzione di un rebuild completo della base di dati, il che implica una cancellazione di tutti i dati presenti nel database MySQL e una nuova configurazione partendo dai file di configurazione, primo tra tutti il file "init.conf".

   Quando si fa un rebuild si provvede a recuperare nuovamente gli attributi associati ad ogni progetto, i parametri per l'inferenza iniziali, le specifiche di rischi, azioni di mitigazione e azioni di recovery disponibili al sistema nel file "specs.conf" e infine i progetti che costituiscono la base di conoscenza iniziale presenti nel file "projects.conf".

6. *Configurazione dei parametri per l'inferenza*

   Alcuni utili parametri per l'inferenza (ad esempio il valore di K), possono essere cambiati durante l'esecuzione del software senza fare un rebuild, per questo è prevista un'interfaccia accessibile tramite un controllo della finestra principale che permette di variare il valore di tali parametri presenti nella base di dati. Ricordiamo nuovamente che i parametri di cui stiamo parlando sono:

   - *K* = costante di contingency
   - *MAX_PROJ_PER_GROUP* = numero massimo di progetti estratti per ogni gruppo
   - *MIN_RI_THRESHOLD* = minimo indice di rischio per un rischio rilevante
   - *MAX_RISKS_PER_PROJ* = massimo numero di rischi che si vuole considerare in ogni progetto
   - *MIN_MID_EFFECTIVENESS* = minima efficacia media percentuale per una buona azione di mitigazione
   - *MAX_MID_INEFFECTIVENESS* = massima inefficacia media percentuale per una azione di mitigazione che non sia troppo dannosa
   - *MOD_THRESHOLD* = soglia di memorizzazione

I punti da 7 a 24 descrivono operazioni possibili su progetti aperti che corrispondono a modifiche per la correzione di incompletezze o errori riscontrati nel risultato del sistema CBR di risk management.

Ogni azione comandata dall'utente in questa categoria causa *l'incremento del campo* ***revision*** all'interno del progetto che subisce la modifica.

7. *Variazione del valore di costo di un'azione di mitigazione non "closed"*

   Ogni azione di mitigazione presente nel nuovo progetto può subire una variazione al suo valore di costo, modifiche da parte dell'utente vengono fatte sempre lavorando in euro indipendentemente dal valore di rel_act. Non sono consentite ulteriori modifiche una volta che l'azione viene chiusa.

8. *Variazione del valore di efficacia in termini di probabilità e impatto di un'azione di mitigazione non "closed"*
   L'utente può allora lavorare sui legami che collegano azioni di mitigazione e rischi, cambiando l'efficacia che un'azione ha su ogni rischio (lavorando in questo caso con valori percentuali) sia in termini di impatto che di probabilità. Questo fintanto che l'azione di mitigazione non viene chiusa, a tal punto i valori sono congelati.

9. *Variazione del valore di efficacia in termini di probabilità e impatto di un'azione di mitigazione "closed" collegata ad un rischio in uno dei due stati closed ma tale che i suoi effetti non compaiono nello storico del rischio*
   Quando un'azione viene chiusa pur essendo collegata ad un rischio chiuso, il sistema sa che in realtà l'azione non può aver avuto un effetto sul rischio (in quanto il rischio si è già verificato o è diventato non significativo). Anche se l'azione è chiusa il legame viene sempre utilizzato dal sistema per l'inferenza come se l'azione fosse non "closed", esso funge da deposito di informazione utile che ci è stata fornita dall'utente. Essendo dunque un deposito, il sistema consente all'utente di cambiare i valori associati anche post-chiusura.

10. *Variazione del valore di contingency di un'azione di recovery relativo ad rischio in stato non "closed_verified"*
    All'utente è concesso di operare sui legami che collegano azioni di recovery e rischi, cambiando la contingency di un'azione di recovery relativa ad ogni rischio non "closed_verified", ossia il valore di risk_contingency (in questo caso l'utente lavora con valori monetari espressi in euro, il sistema effettua la traduzione nel caso in cui rel_risk sia vero dividendo il numero inserito per il valore del progetto).
    Una modifica a posteriori del verificarsi del rischio è vietata in quanto il sistema assume che l'azione di recovery sia stata effettuata (col valore di contingency associato) al momento in cui gli effetti del rischio si sono verificati. Notare che invece sono consentite modifiche quando lo stato del rischio è "closed_unverified", il che significa che il mantenimento del legame è a puro scopo di archiviazione di informazioni per le future inferenze.

11. *Eliminazione di interi rischi non "closed_verified"*
    In questo caso si eliminano le azioni di mitigazione/recovery presenti nel progetto se utilizzate solo per uno dei rischi eliminati e (per quanto riguarda le azioni di mitigazione) se in stato non chiuso. Ovviamente si prevede anche a cancellare ogni legame che il rischio appena eliminato aveva con azioni di mitigazione o di recovery, cancellando i dati contenuti nelle strutture di legame. Le azioni in stato chiuso collegate al solo rischio eliminato vengono invece mantenute (a meno che l'utente non esegua cancellazione esplicita) per tenere memoria del costo che l'azione ha avuto.

12. *Eliminazione di azioni di mitigazione non "closed" collegate a rischi*

    È possibile in questo caso che uno o più rischi cambino strategia di gestione, in particolare:

    - un rischio che era gestito solamente con tale azione di mitigazione (niente recovery) subisce una variazione della sua strategia di gestione in "spostamento".
    - un rischio che era gestito invece con una combinazione di mitigazione e accettazione ma dove l'unica azione di mitigazione era quella che si è eliminata subisce una variazione della sua strategia di gestione in "accettazione".

    Si eliminano allora ovviamente tutte le connessioni che collegavano l'azione appena eliminata con dei rischi. Può anche cambiare lo stato di uno o più rischi, come descritto in sezione 2.

13. *Eliminazione di azioni di mitigazione non collegate a rischi*

    È consentita l'eliminazione di azioni di mitigazione non collegate a rischi indipendentemente dal loro stato.

14. *Eliminazione di azioni di mitigazione "closed" collegate solamente a rischi in uno stato closed e che non hanno avuto effetto su nessuno dei rischi (chiuse dopo la chiusura dei rischi)*

    È possibile in questo caso che uno o più rischi cambino strategia di gestione, in particolare:

    - un rischio che era gestito solamente con tale azione di mitigazione (niente recovery) subisce una variazione della sua strategia di gestione in "spostamento".
    - un rischio che era gestito invece con una combinazione di mitigazione e accettazione ma dove l'unica azione di mitigazione era quella che si è eliminata subisce una variazione della sua strategia di gestione in "accettazione".

    Si eliminano allora ovviamente tutte le connessioni che collegavano l'azione appena eliminata con dei rischi.

15. *Eliminazione di azioni di recovery collegate a rischi non "closed_verified"*

    È possibile in questo caso che uno o più rischi cambino strategia di gestione, in particolare:

    - un rischio che era gestito solamente con tale azione di recovery (niente mitigazione) subisce una variazione della sua strategia di gestione in "spostamento".
    - un rischio che era gestito invece con una combinazione di mitigazione e accettazione subisce una variazione della sua strategia in "mitigazione" (infatti ricordiamo che al più una unica azione di recovery può essere utilizzata per gestire un rischio con strategia di accettazione).

    Si eliminano allora ovviamente tutte le connessioni che collegavano l'azione appena eliminata con dei rischi. Se alcuni rischi associati all'azione di recovery sono chiusi e verificati, l'operazione viene impedita in quanto si assume che l'effetto dell'azione che si vuole eliminare ci sia già stato.

16. *Eliminazione di legami rischio - azione di mitigazione non "closed"*

    Per ogni azione di mitigazione prevista nel progetto non "closed", l'utente può modificare i rischi che tale azione è in grado di mitigare rimuovendo dei legami. L'eliminazione di un legame comporta che il rischio coinvolto può cambiare la sua strategia di gestione, in particolare:

    - Se il rischio era gestito solamente con tale azione di mitigazione (niente recovery) subisce una variazione della sua strategia di gestione in "spostamento".
    - Se il rischio era gestito invece con una combinazione di mitigazione e accettazione ma dove l'unica azione di mitigazione era quella che si è eliminata subisce una variazione della sua strategia di gestione in "accettazione". Può anche cambiare lo stato del rischio, come descritto in sezione 2.

17. *Eliminazione di legami rischio in uno stato closed - azione di mitigazione "closed", fattibile solo se l'azione non ha avuto effetto sul rischio in quanto chiusa a posteriori della chiusura del rischio*

    Un'azione che è stata chiusa può non avere avuto effetti se collegata a rischi che erano già chiusi al momento in cui l'azione è stata chiusa. In questo caso viene consentita la cancellazione da parte dell'utente

18. *Eliminazione di legami rischio non "closed_verified" - azione di recovery*

    Per ogni azione di recovery prevista nel nuovo progetto, l'utente può modificare i rischi non "closed_verified" che tale azione è in grado di recuperare rimuovendo dei legami. L'eliminazione di un legame comporta che il rischio coinvolto cambierà la sua strategia di gestione, in particolare:

    - Se il rischio era gestito solamente con tale azione di recovery (niente mitigazione) subisce una variazione della sua strategia di gestione in "spostamento".
    - Se il rischio era gestito invece con una combinazione di mitigazione e accettazione subisce una variazione della sua strategia in "mitigazione" (infatti ricordiamo che al più una unica azione di recovery può essere utilizzata per gestire un rischio con strategia di accettazione).

    Se l'azione di recovery coinvolta rimane a questo punto isolata da ogni rischio si provvede ad eliminare l'azione stessa come descritto sopra (infatti non è prevista per nessun rischio). Se il rischio a cui l'azione è legata è chiuso e verificato, l'operazione viene impedita in quanto si assume che l'effetto dell'azione ci sia già stato e il legame è significativo. Può anche cambiare lo stato del rischio, come descritto in sezione 2.

19. *Aggiunta di interi rischi*

L'utente può aggiungere rischi prelevandoli esclusivamente dalla lista dei rischi già noti, ossia scegliendo tra i rischi di cui sono già note le specifiche.

Per ogni rischio aggiunto è possibile scegliere quali azioni tra le presenti hanno un effetto di mitigazione e di recovery per tale rischio. A seconda delle scelte dell'utente, si provvederà ad aggiornare opportunamente (come descritto in sezione 2) il campo strategia di gestione, e di conseguenza lo stato.

Se vengono scelte azioni di mitigazione già chiuse nell'ambito del progetto, si va a modificare lo storico del nuovo rischio. La corretta interpretazione è la seguente: il rischio era presente nel progetto dall'origine (infatti si assume che i progetti non cambino nel tempo) quindi se un'azione doveva avere un effetto su tale rischio allora l'ha già avuto ed è giusto aggiornare lo storico del rischio. Quando si aggiunge un rischio si inserisce il corretto valore all'interno del campo *revision* del rischio stesso, visto in sezione 2.

20. *Aggiunta di azioni di mitigazione*

L'utente può aggiungere azioni di mitigazione prelevandole esclusivamente dalla lista delle azioni di mitigazione già note, ossia scegliendo tra le azioni di mitigazione di cui sono già note le specifiche.

Per ogni azione aggiunta, l'utente deve decidere a quali rischi presenti nel progetto tale azione è associata in quanto ha o avrebbe avuto su di essi un effetto di mitigazione. L'utente deve inoltre specificare il costo per l'azione di mitigazione nel progetto in esame e l'efficacia di tale azione su ogni rischio ad essa legato in termini di probabilità e impatto sul rischio.

I rischi a cui l'azione è stata associata possono potenzialmente cambiare strategia di gestione e stato, ad aggiornare tali campi ci pensa il sistema secondo la logica descritta in sezione 2.

L'azione viene impostata in stato "planned".

21. *Aggiunta di azioni di recovery*

L'utente può aggiungere azioni di recovery prelevandole esclusivamente dalla lista delle azioni di recovery già note, ossia scegliendo tra le azioni di recovery di cui sono già note le specifiche.

Per ogni azione aggiunta, l'utente deve decidere a quali rischi non "closed_verified" presenti nel progetto tale azione è associata in quanto viene usata per il recupero degli stessi. L'utente deve in questo caso specificare quale è la frazione di contingency dell'azione di recovery destinata a ciascun rischio stabilendo i valori dei campi risk_contingency presenti in ogni connessione. I valori dei campi original_RI non vengono specificati dall'utente in quanto calcolati dal sistema come specificato precedentemente e dunque utilizzando i valori attuali di impatto e probabilità dei rischi quando questi sono in stato non "closed_unverified".

Se i rischi sono invece in stato "closed_unverified" si utilizzano i valori iniziali di impatto e probabilità per il calcolo di original_RI, ad indicare che un'azione di recovery sarebbe stata opportuna fin dall'inizio.

I rischi a cui l'azione è stata associata possono potenzialmente cambiare strategia di gestione e stato, ad aggiornare tali campi ci pensa il sistema secondo la logica descritta in sezione 2.

22. *Aggiunta di legami rischio - azione di mitigazione non "closed"*

Per ogni azione di mitigazione non chiusa prevista nel progetto, l'utente può modificare i rischi che tale azione è in grado di mitigare, aggiungendo o rimuovendo dei legami.

Nel caso in cui l'utente voglia creare dei nuovi legami, egli è comunque obbligato a specificare l'efficacia dell'azione in questione in termini di probabilità e impatto sul rischio al quale questa azione si sta legando (come valori percentuali).

23. *Aggiunta di legami rischio in uno stato closed - azione di mitigazione "closed"*

Le azioni di mitigazione chiuse possono essere collegate dall'utente ad un rischio in uno stato closed, ad indicare (a posteriori della chiusura del rischio) che l'azione poteva avere in realtà un effetto di qualche tipo su tale rischio. L'utente deve specificare in questo caso dei valori di efficacia prospettici (che il sistema considera col giusto peso in fase di inferenza).

24. *Aggiunta di legami rischio non "closed_verified" - azione di recovery*

Per ogni azione di recovery prevista nel progetto, l'utente può modificare i rischi che tale azione deve recuperare, aggiungendo o rimuovendo dei legami.

Nel caso in cui l'utente comandi una simile operazione, è però obbligato anche a definire il nuovo valore per la contingency di quella azione destinata a recuperare il rischio a cui è stata legata (valore di risk_contingency). Il valore del campo original_RI non viene specificato dall'utente in quanto calcolato dal sistema come specificato precedentemente e dunque utilizzando i valori attuali di impatto e probabilità del rischio quando questo è in stato non "closed_unverified".

Notiamo che si possono aggiungere legami anche con rischi chiusi ma non verificati, con questa operazione l'utente indica che una recovery sarebbe stata auspicabile per il rischio chiuso (anche se non era stata prevista all'inizio). In questo caso per il calcolo di original_RI si utilizzano i valori originali per impatto e probabilità del rischio.

6 Case Study: Traduzione dei dati e configurazione del sistema

In sezione 3 è stata descritta la struttura della base di dati dove il sistema immagazzina tutte le informazioni che ha a disposizione. È ovvio che quando il CBR Risk Management System verrà applicato in azienda, esso dovrà venire opportunamente configurato tramite i file di configurazione. Un problema che ci può essere in questa fase è che alcune informazioni richieste dal sistema possono non essere normalmente gestite dalla politica aziendale, oppure può essere il caso che alcune informazioni a disposizione dall'azienda non siano considerate rilevanti dal sistema e non siano previste per il suo funzionamento.

Come esempio di configurazione opportuna riportiamo allora un *Case Study* dove un database aziendale viene tradotto in una opportuna configurazione del CBR Risk Management System.

La descrizione dei dati viene fornita tramite 5 tabelle:

1. **Progetto**
   - **ID Programma**: stringa che rappresenta l'identificativo del programma
   - **Reparto**: enumerato nell'insieme {1,2,3}
   - **Classe di rischio**: enumerato nell'insieme {alto, medio, basso}
   - **Valore economico**: numero reale che rappresenta il valore economico del progetto
   - **Durata contratto**: numero intero, anni di durata del contratto
   - **Oggetto**: oggetto della fornitura
   - **Cliente**: nome del cliente
   - Variabili che influenzano il **livello di rischio**, tutti enumerati con valori nell'insieme {0,1,2,3}, organizzate in gruppi:
     - PAESE
       - RP1
       - RP2
       - RP3
     - MERCATO CLIENTE
       - RMC1
       - RMC2
       - RMC3
     - CONTRATTO
       - RC1
       - RC2
       - RC3

- COMPOSIZIONE PARTNERSHIP
  - RCP1
  - RCP2
  - RCP3
- INGEGNERIA
  - RI1
  - RI2
  - RI3
- APPROVVIGIONAMENTO
  - RA1
  - RA2
  - RA3
- FABBRICAZIONE
  - RF1
  - RF2
  - RF3
- MONTAGGIO
  - RM1
  - RM2
  - RM3
- AVVIAMENTO
  - RAV1
  - RAV2
  - RAV3

- Variabili influenzanti **l'impatto strategico**, ognuna è un enumerato che può assumere valori nell'insieme {0,1,2,3}
  - importanza del mercato (IM)
  - interesse nell'acquisire il cliente (IC)
  - grado di innovazione del prodotto (IP)
  - incidenza sull'immagine aziendale (IA)
  - interesse nella partecipazione alla partnership (IPP)

2. Rischio

- **Id programma**: identificativo del programma (progetto) a cui il rischio fa riferimento
- **Id rischio**: codice identificativo del rischio (univoco all'interno della base di dati)
- **Cod Checklist**: codice del rischio (relativo alla checklist aziendale)
- **Stato**: stato del rischio (mitigazione, monitoraggio, chiuso)
- **Categoria**: categoria a cui il rischio appartiene
- **R_Ver**: valore che indica se il rischio si è o meno verificato come segue:
  0 - rischio aperto
  1 - rischio chiuso e verificato
  2 - rischio chiuso ma non verificato

- **Numero Revisione**: numero indicante la revisione attuale del progetto
- **Descrizione del rischio**: stringa che descrive il rischio
- **Contingency**: numero reale che rappresenta la contingency accantonata
- **Probabilità iniziale**: probabilità in revisione 0
- **Impatto**: impatto in revisione 0, è un indice impatto (vedi oltre)

3. **Cause ed effetti dei rischi**
   - **Id_rischio**: codice identificativo del rischio (come sopra)
   - **Cause**: stringa che descrive le cause del rischio
   - **Effetti**: stringa che descrive gli effetti del rischio

4. **Azioni legate ai rischi**
   - **Id_rischio**: codice identificativo del rischio (come sopra)
   - **Id_azione**: identificativo dell'azione nell'ambito della gestione di un rischio
   - **Cod_check:** codice checklist dell'azione nella checklist delle azioni aziendale
   - **Tipo**: mitigazione o recovery
   - **Stato**: stato in cui si trova l'azione (pianificata,in progress, chiusa, di back-up)
   - **Descrizione**: descrizione dell'azione
   - **Revisione**: revisione in cui l'effetto di un'azione di mitigazione è stato inserito nello storico (ossia l'azione è stata eseguita divenendo chiusa) se diverso da 0
   - **Intensità**: valore tra -10 e 10 che indica l'intensità dell'effetto dell'azione su probabilità e impatto

5. **Storico dei rischi**
   - **Id_rischio**: codice identificativo del rischio (univoco all'interno della base di dati)
   - **Numero revisione**: numero indicante la revisione in cui il rischio ha cambiato probabilità e impatto per un qualche effetto sconosciuto al sistema CBR.
   - **Probabilità**: numero in [0,100] che indica la probabilità ce il rischio si verifichi
   - **Indice impatto**: enumerato in {1,2,3,4,5} dove 1 rappresenta il minimo impatto e 5 il massimo.

6. **Ck_mitigazione**
   - **Cod Checklist**: codice che identifica l'azione di mitigazione nella checklist aziendale
   - **Descrizione**: descrizione generica dell'azione di mitigazione nella checklist aziendale

7. **Ck_recovery**
   - **Cod Checklist**: codice che identifica l'azione di recovery nella checklist aziendale
   - **Descrizione**: descrizione generica dell'azione di recovery nella checklist aziendale

Cominciamo allora ad esplicitare come tutti questi campi possono venir tradotti nei campi delle strutture presenti nel sistema CBR. Bisogna considerare che comunque le scelte seguenti dovrebbero essere fatte dall'esperto aziendale che si occupa del risk management. Quando costui configura il sistema è a conoscenza di alcuni parametri ignoti in questo documento, quanto segue è invece solamente una possibile soluzione scritta a scopo esemplificativo.

Intanto possiamo dire che, essendo gli attributi presenti nei progetti gestiti dal sistema CBR completamente arbitrari e configurabili, l'unica cosa da fare è scegliere come organizzare i vari attributi in gruppi.

- Gruppo **generic**: ID programma, reparto, classe di rischio, valore economico, durata contratto, oggetto, cliente.
- Un gruppo per ogni gruppo di variabili influenzanti il livello di rischio, avente lo stesso nome del gruppo fornitoci dall'azienda e le stesse variabili come membri.
- Gruppo **strategic**: IM, IC, IP, IA, IPP.

Facciamo inoltre alcune osservazioni per gestire in modo consistente i dati forniti dall'azienda qualora risultino incompleti o non trattabili dal sistema.

- ✓ Per alcuni rischi forniti dall'azienda nell'ambito di particolari programmi, non sono stati inseriti i valori di indice di impatto e probabilità iniziale, per tali casi si assume:
  - Probabilità iniziale = 30
  - Indice impatto iniziale = 3

  Che sono approssimativamente dei valori medi per le informazioni a disposizione.
- ✓ Talvolta ad un singolo progetto sono associati diversi rischi aventi lo stesso codice di checklist, questa situazione non è gestibile dal sistema in quanto si è assunto che un rischio (in termini generali) sia identificato dal codice checklist e possa presentarsi una sola volta nell'ambito di un progetto (non avrebbe senso avere più volte lo stesso rischio).

  Si provvede dunque ad effettuare una *fusione* di rischi con lo stesso codice checklist nell'ambito di un solo progetto, facendolo:
  - Si sommano le contingency
  - Si mediano le loro probabilità e i loro impatti
  - Lo stato del rischio risultante non è chiuso a meno che chiusi non siano tutti i rischi che vengono fusi tra loro
- ✓ In azienda si prevede la situazione in cui un rischio, pur essendo gestito solamente tramite azioni di recovery, è in stato "monitoring". Si ricorda invece il differente significato che è stato associato allo stato "monitoring" nel CBR Risk Management System.

  Se il rischio è gestito solamente con azioni di recovery la configurazione del sistema avviene con lo stato "recovery".
- ✓ Gli impatti iniziali dei rischi sono stati forniti dall'azienda come *indici di impatto*, abbiamo però già detto che il CBR Risk Management System lavora con *costi potenziali di impatto*.

  In questa ottica allora si è deciso di assegnare ad ogni rischio rel_risk pari a true per avere impatti percentuali relativi al valore del progetto, e di effettuare una traduzione dell'indice impatto fornito in un costo potenziale come segue:

- indice impatto = 1 → costo potenziale impatto = 4% del valore del progetto
- indice impatto = 2 → costo potenziale impatto = 8% del valore del progetto
- indice impatto = 3 → costo potenziale impatto = 12% del valore del progetto
- indice impatto = 4 → costo potenziale impatto = 16% del valore del progetto
- indice impatto = 5 → costo potenziale impatto = 20% del valore del progetto

✓ Nelle informazioni fornite dall'azienda abbiamo che quando un rischio è gestito tramite azioni di mitigazione lo stato del rischio è sempre mitigation o closed.
La gestione dello stato del rischio nel CBR Risk Management System invece è diversa ed è quella che è stata esaminata in sezione 2, per cui si provvede a modificare l'informazione fornitaci in fase di configurazione in dipendenza delle azioni che sono state previste per ogni rischio (anche quelle di recovery).

✓ Se non ci sono azioni di alcun genere associate ad un rischio si è assunto che la strategia per quel rischio fosse "transfer" e come stato si è imposto "monitoring".

✓ Lo stato di un programma si è assunto "closed" qualora tutti i rischi ad esso associati siano chiusi, altrimenti lo stato è "active".

✓ In analogia con quanto presente nelle informazioni fornite dall'azienda, la revisione di un'azione di mitigazione nell'ambito di un particolare progetto viene posta a 0 ad indicare che l'azione non è ancora stata eseguita completamente. Un valore diverso da 0 sarà presente solamente quando l'azione viene eseguita e di conseguenza chiusa, in tal caso si memorizza la revisione del progetto al momento della chiusura dell'azione di mitigazione.

✓ Il numero di revisione del progetto è stato individuato come il numero di revisione più alto associato a quel progetto presente nelle tabelle che ci sono state fornite dall'azienda.

✓ Il campo revisione associato ad un rischio nell'ambito di un programma presente nella tabella "Rischio" indica la revisione attuale del programma stesso, noi invece necessitiamo del numero di revisione del programma al momento in cui il rischio è stato aggiunto. Per semplicità noi consideriamo che tutti i rischi siano stati aggiunti inizialmente ai progetti, per cui il numero di revisione risulta 0.

✓ Nelle informazioni fornite dall'azienda si possono scorgere delle azioni di recovery in stato "planned". Nel CBR Risk Management System le azioni di recovery non dispongono di uno stato nell'ambito del progetto, per cui questa informazione è stata ignorata (come se lo stato fosse "back-up").

✓ I valori di intensità delle azioni di mitigazione sono rappresentati come interi tra -10 e 10 nelle informazioni fornite dall'azienda. Il CBR Risk Management System necessita di conoscere i valori di efficacia percentuale sull'impatto e sulla probabilità. Per trattare questi valori in modo consistente, si è deciso semplicemente di moltiplicare per 10 il valore presente nella tabella per ottenere l'efficacia in termini di probabilità e impatto.

*impact_effectiveness = probability_effectiveness = intensità * 10*

✓ Per le azioni per cui non viene riportata l'intensità (perché non chiuse) il sistema CBR ha bisogno di una **efficacia prospettica** sia sulla probabilità che sull'impatto del rischio, si assume dunque in questo caso un valore di default pari al *20%* in entrambi i campi.

- ✓ Vi sono dei casi nelle informazioni forniteci in cui ai rischi è associata una contingency non nulla mentre il valore del progetto è nullo.
  Nel CBR Risk Management System la risk_contingency associata ad un rischio tramite un'azione di recovery, quando rel_risk è a true, è un valore percentuale che indica quanto denaro è necessario spendere per recuperare gli effetti del rischio qualora si verifichi, in relazione al valore del progetto.
  Essendo il valore del progetto nullo (situazione assurda) allora l'impatto del rischio è nullo (valutato in euro). La risk_contingency non può allora essere inserita normalmente. Nel CBR Risk Management System, la risk_contingency è considerata come il denaro necessario per recuperare gli effetti di un rischio. Se un rischio ha impatto nullo (in euro), è ovvio che la risk_contingency avrà valore nullo in euro in quanto il rischio non ha alcun costo e allora non è necessario spendere denaro per recuperare i suoi effetti. Un rischio presente in un progetto a valore nullo (inutile) non può avere degli effetti significativi e dunque ha sempre impatto in euro nullo.
  Una risk_contingency pari a 0 euro, visto che il valore del progetto è nullo, può essere espressa usando qualunque percentuale: si inserisce allora come risk contingency il valore di partenza, calcolato come
  $$VPart = original_RI * K$$
  Facendo questa scelta abbiamo che l'azione di recovery non verrà privilegiata nel processo di inferenza, in quanto non si può sapere quanto il progetto sia effettivamente minacciato dal rischio che si è riuscito a recuperare con l'azione in esame (e dunque quanto buona è stata questa azione).
  In questo modo si perdono i dati di contingency forniti dall'azienda nel caso in cui il valore del progetto sia nullo ma la contingency associata ad un rischio sia non nulla, infatti questi sono ritenuti non validi dal sistema.
- ✓ Si fissano i valori di costo di tutte le azioni di mitigazione pari a 0, infatti questo parametro non è previsto dall'azienda.

Diamo adesso una tabella di corrispondenza per associare ad ogni campo necessario al CBR Risk Management System un valore estratto dai dati forniti dall'azienda.

Valori Forniti	**Campi del CBR**
Project	
ID Programma	company_ID [S]
Reparto	<in features> [E]
Classe di rischio	<in features> [E]
Valore economico	value [N]
Durata contratto	<in features> [N]
Oggetto	service_material [S]
Cliente	customer name [S]

Variabili livello di rischio	<in features> [E]
IM	<in features> [E]
IC	<in features> [E]
IP	<in features> [E]
IA	<in features> [E]
IPP	<in features> [E]
Si prende il massimo per le revision incontrate	revision
Chiuso a meno che non ci siano rischi non chiusi	state
VUOTO	project name [S]
Generato automaticamente	project_ID
Proj_Risk	
Legati ai Cod Checklist in tabella 2, ma ricavati in modo incrementale	risk_ID
Id_rischio	company_risk_ID
Ricavata dalle azioni associate	strategy
State + eventualmente recovery Per la valutazione dello stato si utilizza anche il campo R_Ver in tabella 2	state
Numero revisione	revision
Si traduce l'indice impatto come segue: 1->4% 2->8% 3->12% 4->16% 5->20% del valore del progetto, con rel_risk pari sempre a false. Si inferisce l'ordine di applicazione delle azioni a "Numero Revisione" in tabella 4. L'impatto iniziale si ricava dalla tabella 2, campo "Impatto"	impact
Probabilità Iniziale, come sopra per inferire sullo storico. La probabilità iniziale si ricava dalla tabella 2, campo "probabilità iniziale"	probability
Descrizione del rischio	description
Insieme di cause in tabella 3	causes

Insieme di effetti in tabella 3	effects
R_Ver	La condizione di rischio verificato è codificata nel sistema CBR all'interno dello stato del rischio
<u>Proj_Mit_Bind</u>	
Da "Intensità" in tabella 4, con le osservazioni di cui sopra	Impact/probability effectiveness per ogni azione di mitigazione associata al rischio
<u>Proj_Rec_Bind</u>	
Contingency (l'azienda assume il costo di tutte le azioni di mitigazione pari a 0)	risk_contingency *non* va normalizzato, perché rel_risk è sempre pari a "false"
Automaticamente calcolato al momento della lettura dei file di configurazione a partire da impatto e probabilità del rischio collegato nell'ambito del progetto	original_RI
<u>Proj_Mit_Action</u>	
Cod Checklist in tabella 4	mitigation_ID
0	cost
Descrizione in tabella 4	description
Revisione in tabella 4	revision
Pianificata -> planned In progress -> active Chiusa -> closed	state
<u>Proj_Rec_Action</u>	
Cod Checklist in tabella 4	recovery_ID
Descrizione in tabella 4	description
Stato "di backup"	<non ha stato> Si assume infatti che quando un'azione di recovery viene prevista essa venga sempre eseguita con successo quando il rischio si verifica
<u>Risk</u>	
Legati ai Cod Checklist in tabella 2, ma ricavati in modo incrementale	risk_ID
Cod Checklist in tabella 2	risk_CL_code

Categoria in tabella 2 (unica per una specifica del rischio)	category
Generalizzazione delle descrizioni in tabella 2	description
Unione delle cause in tabella 3 per quello specifico Cod checklist	causes
Unione degli effetti in tabella 3 per quello specifico Cod checklist	effects
FALSE (impatti sempre in euro)	rel_risk
Mit_Action	
Cod Checklist in tabella 6	mitigation_ID
Cod Checklist in tabella 6	action_CL_code
FALSE (si assume, che qualora venga aggiunto un costo, questo vada trattato in euro)	rel_act
Descrizione in tabella 6	description
Rec_Action	
Cod Checklist in tabella 7	recovery_ID
Cod Checklist in tabella 7	action_CL_code
Descrizione in tabella 7	description

I parametri di inferenza sono stati configurati nel seguente modo:

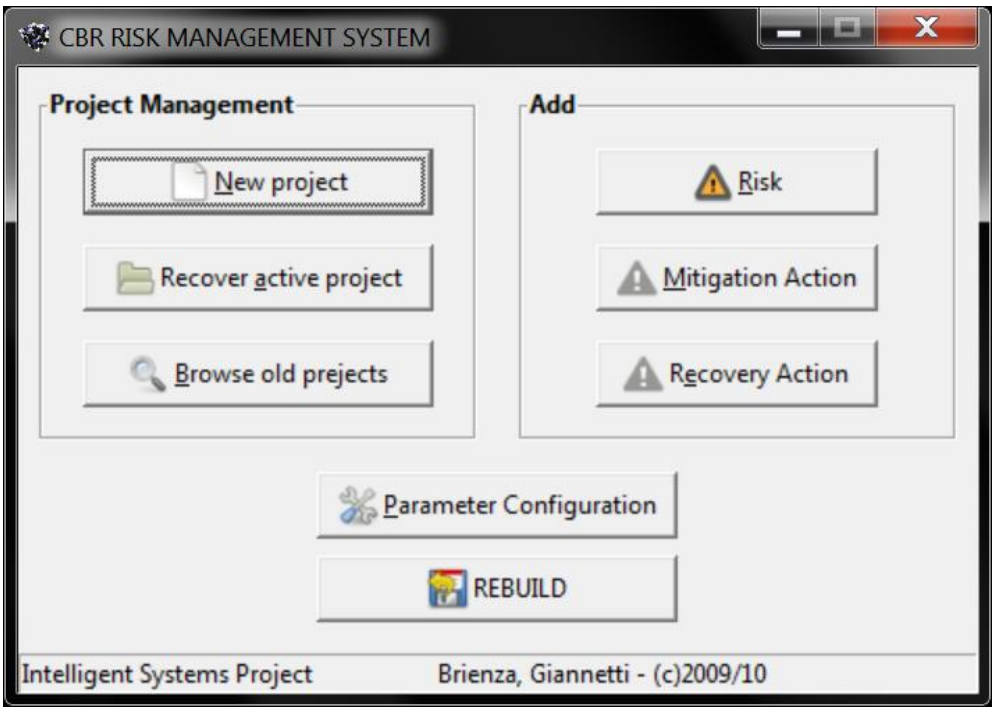
- *K* = 1.0
- *MAX_PROJ_PER_GROUP* = 5
- *MIN_RI_THRESHOLD* = 0.0 (si filtrano i rischi solamente con MAX_RISKS_PER_PROJ)
- *MAX_RISKS_PER_PROJ* = 15
- *MIN_MID_EFFECTIVENESS* = 15.0
- *MAX_MID_INEFFECTIVENESS* = 2.0
- *MOD_THRESHOLD* = 100

Fare riferimento alle sezioni precedenti per chiarire il significato di questi parametri.

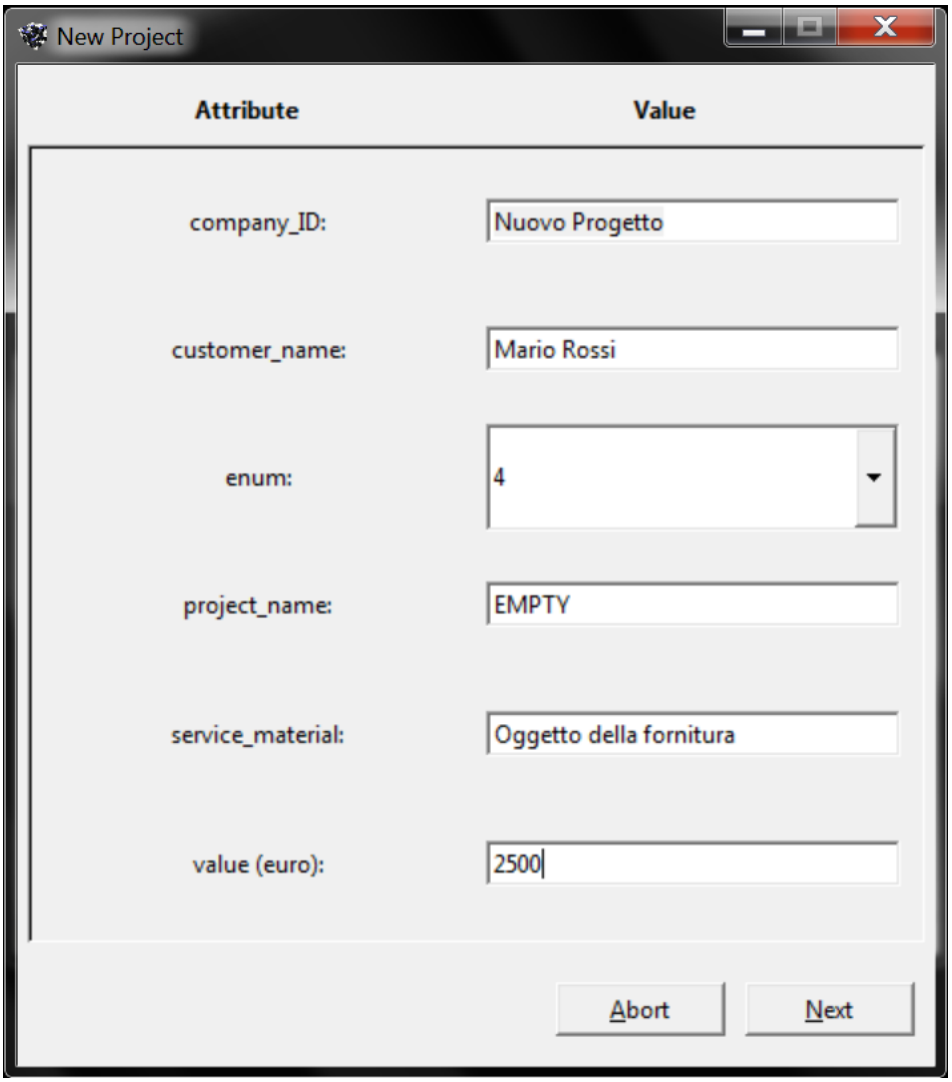
7 Implementazione del sistema

Come ormai si sarà capito leggendo la presente documentazione, il sistema risulta un software molto complesso dove, oltre agli aspetti critici di gestione della procedura di inferenza, abbiamo dialogo con una base di dati e gestione di un'interfaccia grafica non banale.
Per il poco tempo a disposizione dei progettisti, solo una parte delle funzionalità specificate sopra sono state realizzate, di seguito si riporta una lista di cosa è stato fatto:

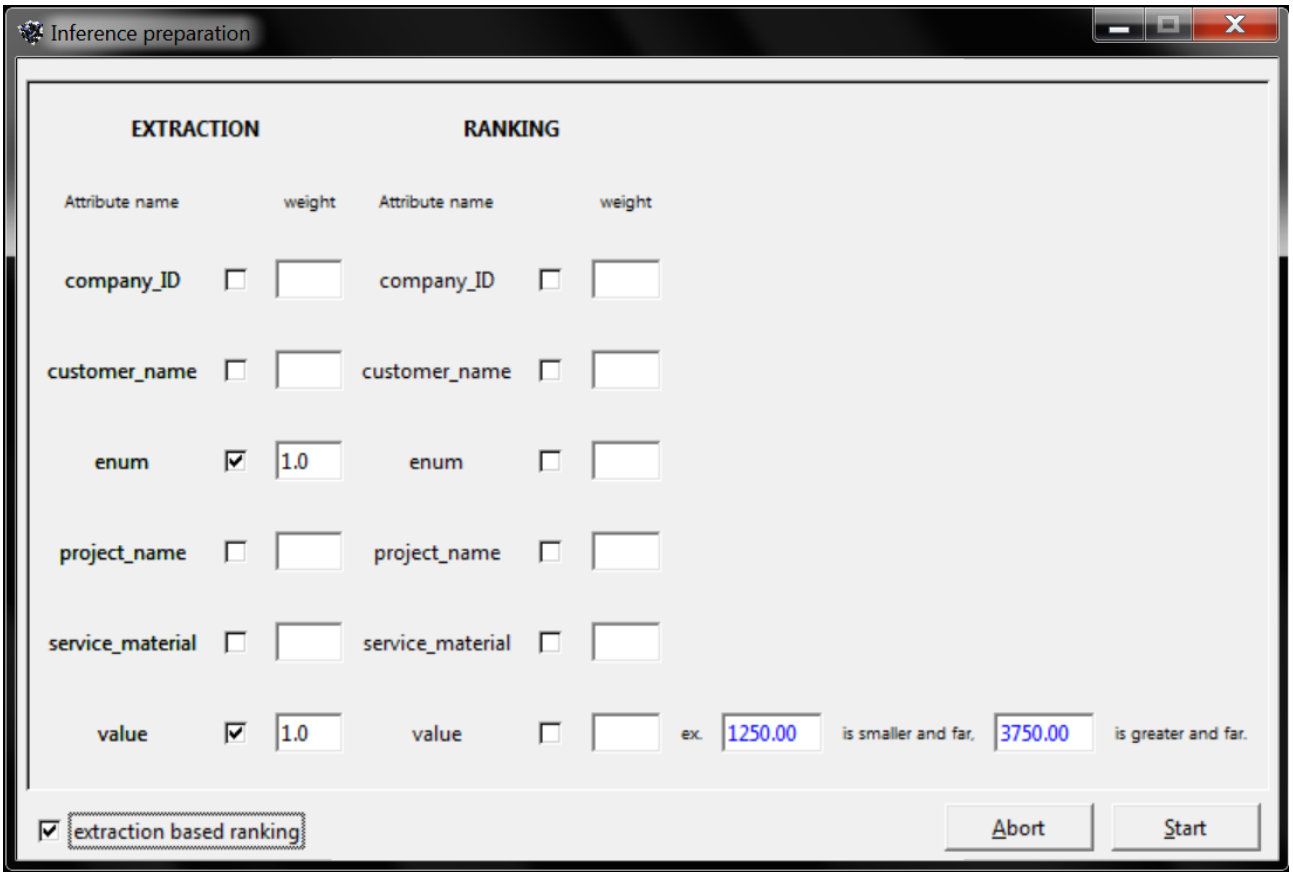
- Implementazione *completa* del **modulo di gestione della conoscenza**, interfacciamento con il database MySQL e parser dei file di configurazione (sottomodulo del modulo di gestione della conoscenza).
- Implementazione *completa* del **modulo di inferenza**, con tutte le caratteristiche spiegate precedentemente.
- Implementazione *parziale* del **modulo di interfaccia**: le possibili schermate da presentare all'utente sono molte, quelle che sono state effettivamente realizzate sono le seguenti:
  - <u>Schermata principale di selezione delle attività da svolgere</u>
    In questa sono presenti diversi pulsanti che dovrebbero consentire di eseguire tutte le funzioni descritte per il modulo di interfaccia, solo un sottoinsieme di queste funzioni sono state implementate e dunque non tutti i pulsanti risultano funzionanti.



- <u>Schermata di inserimento di un nuovo progetto</u>
  Questa consente di inserire i valori per tutte le features del nuovo progetto, alla specifica del valore degli attributi segue immediatamente l'inizio del processo di inferenza.



- <u>Schermata di scelta degli attributi per estrazione e ranking</u>
  Questa consente la specifica degli attributi da utilizzare per estrazione e ranking (tramite checkbox) e i relativi pesi (tramite campi di testo). Inoltre vi è l'opzione per effettuare extraction based ranking.

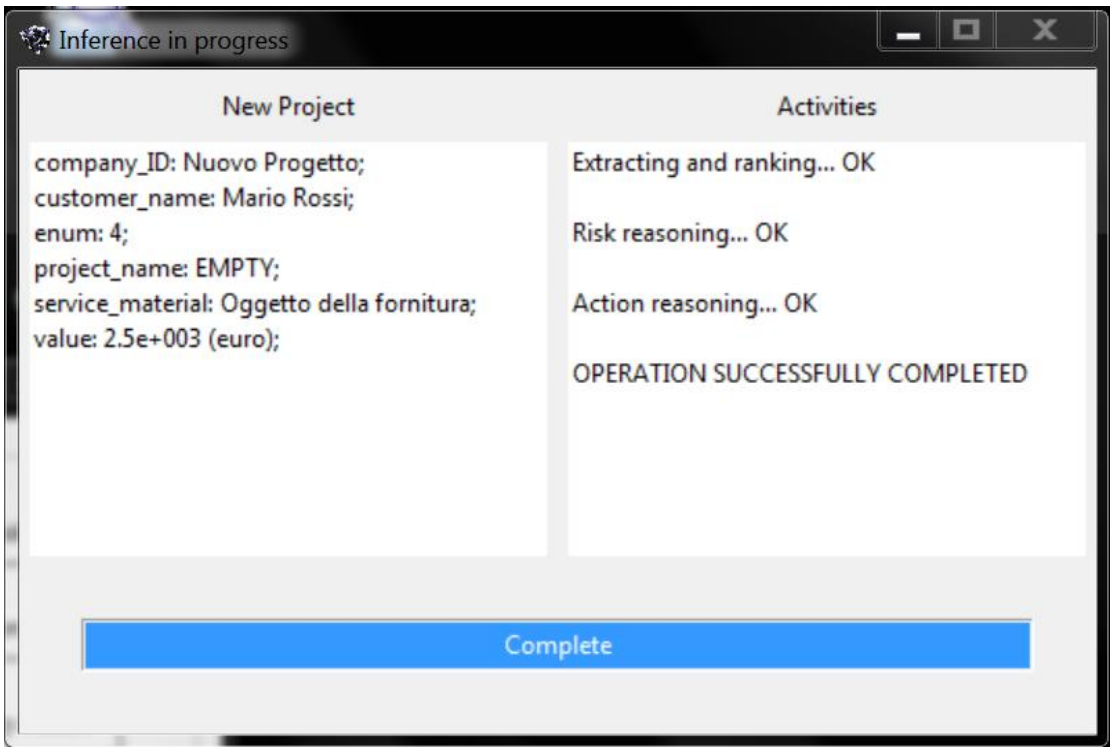


- <u>Schermata di avanzamento del processo di inferenza</u>

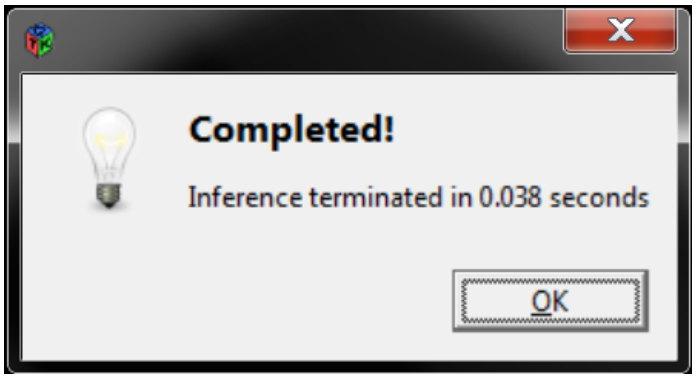
  Ricordando la sezione 4, si ha che un'interazione con l'utente è richiesta per l'identificazione dei rischi e delle azioni di mitigazione, questa dovrebbe avvenire tramite barre di scorrimento, in modo molto intuitivo.

  Per questioni di mancanza di tempo, queste interazioni non sono state implementate e si ricorre al seguente comportamento:

  - Per i rischi non si effettua un ulteriore filtraggio, come se la funzionalità non fosse mai stata prevista;
  - Per le azioni di mitigazione si ricorre al valore di default per il filtraggio, visto in sezione 4 (peso normalizzato pari al 15%).



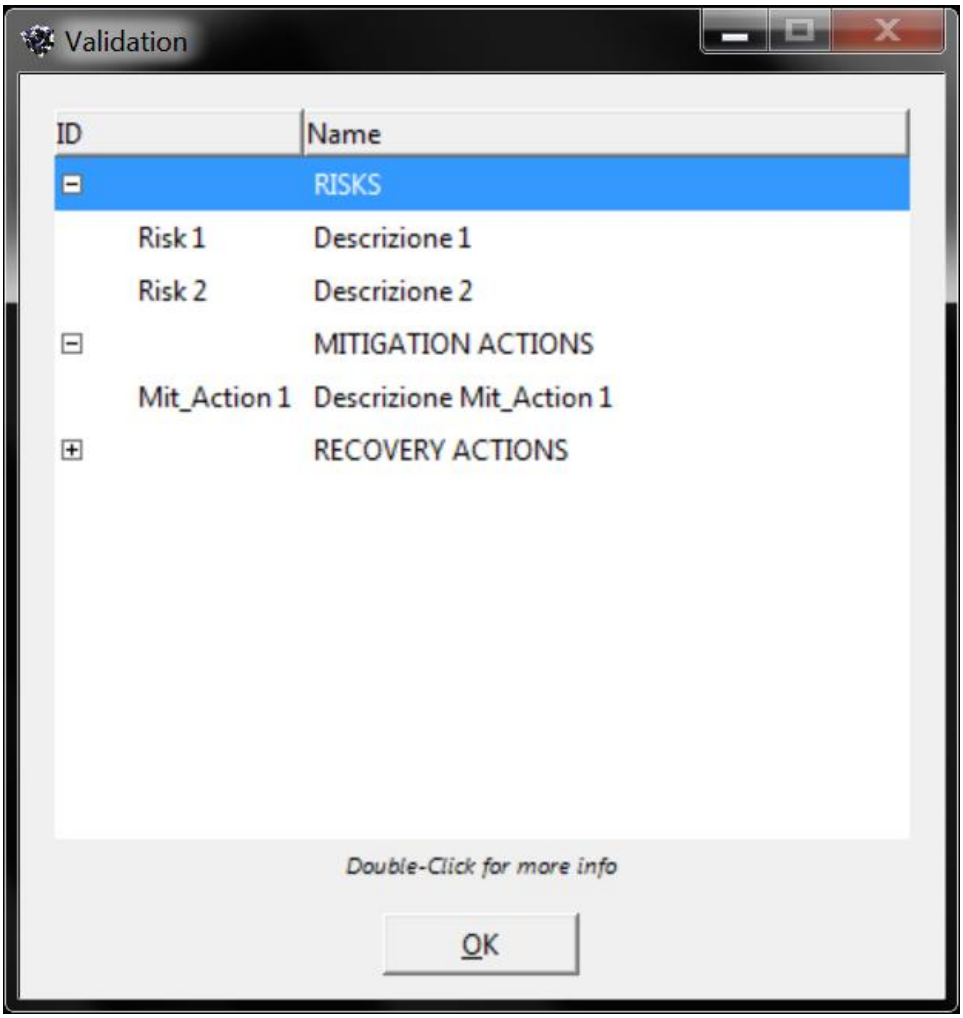
Al termine dell'inferenza una finestra comunica il tempo impiegato per completare il processo:



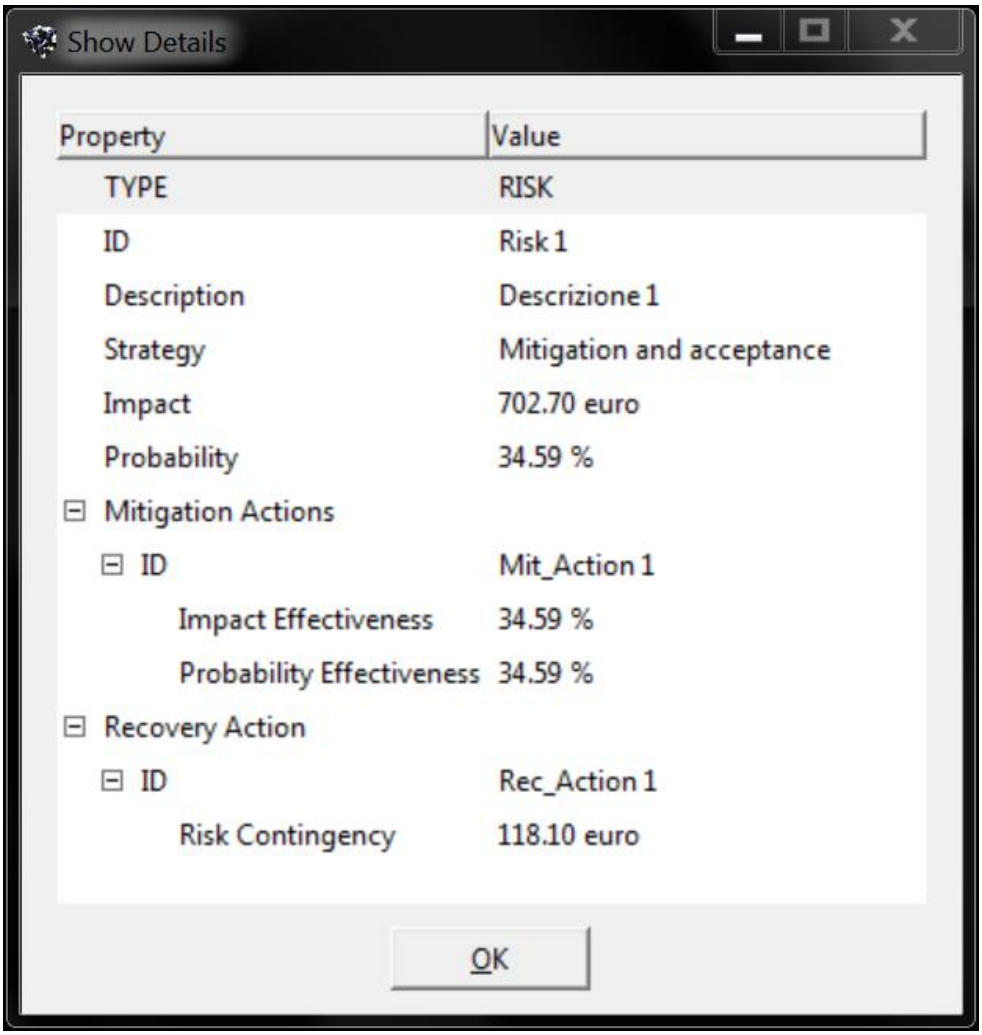
- <u>Schermata di validazione dei risultati dell'inferenza</u>

  Ricordando la sezione 4, è ovvio che questa schermata dovrebbe non solo presentare i dati all'utente ma anche fornire le funzionalità per alterare profondamente il risultato del processo di inferenza.

  Per questioni di mancanza di tempo, questa schermata si occuperà soltanto di presentare il risultato all'utente finale.



Facendo doppio click su una delle voci, viene mostrata una schermata che contiene tutti i dettagli dell'oggetto selezionato, ad esempio per il primo rischio:



Nel caso in cui il database del sistema non sia presente o comunque in stato valido, al primo avvio del sistema viene chiesto all'utente se si vuole fare adesso la costruzione dello stesso. In un qualunque momento l'utente può comandare la ricostruzione dell'intero database a partire dai file di configurazione usando l'apposto comando REBUILD nella schermata principale.

L'intera GUI del sistema è stata pensata per essere utilizzata in qualunque ambito applicativo, pertanto:

- Essa è interamente in lingua inglese (eccetto ovviamente i nomi degli attributi specificati dall'utente, che possono essere specificati in qualsiasi lingua);
- I nomi degli attributi (features) di default del sistema CBR sono quelli visti nella presente documentazione, anche se in disaccordo con i nomi presenti nel case study aziendale preso in considerazione.