versione 0.3
19 dicembre 2007
Giovanni L. Ciampaglia

Wikimodel

Questo documento fornisce le definizioni per il modello matematico di Wikipedia, l'enciclopedia libera della Wikimedia foundation.

Definition (Argomento enciclopedico[1])**.** $\varepsilon \in \mathbb{E}$ denota[2] un argomento enciclopedico. Più precisamente, $\mathbb{E}$ è un insieme finito di etichette.[3]

Gli utenti di Wikipedia possono essere modellizzati come agenti cognitivi raggruppati in popolazioni opportune. Il concetto di popolazione verrà introdotto più avanti, al fine di ridurre il numero di gradi di libertà del sistema.

Definition (Agente)**.** $a \in Ag$ denota un utente[4] ovvero *agente* di Wikipedia. L'insieme Ag è finito.

Definition (Tratto)**.** $\tau \in \mathcal{T}$ denota un tratto semantico. L'insieme $\mathcal{T}$ è un insieme di etichette. È sempre possibile considerare un ordinamento totale[5] su $\mathcal{T}$.

Note. L'insieme $\mathcal{T}$ andrebbe pensato allo stesso modo dell'insieme di valori di una variabile linguistica, nozione proveniente dalla teoria del ragionamento approssimato fuzzy. Questo vuol dire che possono esistere regole grammaticali per costruire un tratto, ma non è nei miei scopi includere tali dettagli in questo modello di Wikipedia.

Sarebbe ingenuo pensare di modellizzare la dinamica dello sviluppo di Wikipedia prendendo in considerazione tutti i contenuti che possono essere immessi in una enciclopedia ipertestuale. Resta il problema di affrontare la differenza tra pagine più o meno importanti, argomenti più o meno generali e via discorrendo. In altre parole, è necessario tenere conto delle relazioni semantiche che caratterizzano ciascun argomento enciclopedico. In questo modello di Wikipedia non verrà data una definizione esplicita di "argomento enciclopedico"; al contrario, un argomento è implicitamente definito in relazione agli altri argomenti presenti nell'enciclopedia. Questo approccio è facilmente realizzabile considerando un grafo delle relazioni semantiche tra gli elementi di $\mathbb{E}$.

Definition (Grafo di $\mathbb{E}$)**.** Dato l'insieme $\mathbb{E}$ esistono un certo numero di collegamenti tra i suoi elementi che possiamo rappresentare con un grafo[6] diretto $G_{\mathbb{E}} = (V(\mathbb{E}), E(\mathbb{E}))$;

[1]Nel modello per un generico sistema sociale questo concetto prende il nome di "contenuto" ovvero "merce" del sistema.

[2]In questo documento, il simbolo $\mathbb{E}$ indica l'insieme della conoscenza enciclopedica (dal Greco *epistémé*); per evitare confusione con il simbolo E, universalmente usato per indicare l'insieme degli archi di un grafo, ho usato la stessa notazione dell'insieme dei numeri reali $\mathbb{R}$.

[3]In generale, $\mathbb{E} = \mathbb{E}(t)$, dato che è naturale pensare la somma della conoscenza umana come un'entità in costante crescita. Tuttavia, questa dinamica è certamente lenta rispetto a quella dell'enciclopedia stessa, per cui nel resto del documento considererò $\mathbb{E}$ come fisso.

[4]Un utente corrisponde, dal punto di vista tecnico, ad un account che interagisce con i server di Wikipedia; poiché molti contributi avvengono in forma anonima, includiamo in questa definizione anche gli account temporanei non registrati che il software MediaWiki identifica con l'indirizzo IP pubblico di provenienza.

[5]Se $\mathcal{T}$ è finito e numerabile, basta fissare una permutazione dei suoi elementi; altrimenti, $\mathcal{T}$ deve essere equipaggiato con una metrica.

[6]In Wikipedia questi collegamenti semantici vengono esplicitati sotto forma di link ipertestuali tra le pagine delle voci enciclopediche.

l'insieme dei nodi è $\mathbb{E}$ e l'insieme degli archi è un sottoinsieme del prodotto cartesiano di $\mathbb{E}$, in simboli $E(\mathbb{E}) \subseteq \mathbb{E} \times \mathbb{E}$. Per ragioni che saranno chiare in seguito, l'insieme degli archi è dinamico: $E(\mathbb{E}) = E_t(\mathbb{E})$

Gli utenti fruiscono dei contenuti dell'enciclopedia in modo attivo, all'occorrenza modificandoli se questi non risultano soddisfacenti; ciò è in larga parte dovuto alle loro valutazioni soggettive. La seguente definizione modellizza dunque le varie opinioni di un agente in merito ad un argomento.

Definition (Punto di vista di un Agente)**.** Sull'insieme degli agenti Ag e per ogni istante di tempo $t \in \mathbb{R}^+$ sono definite una funzione insiemistica τ:

$$\tau : \mathbb{E} \times Ag \times \mathbb{R}^+ \to \mathscr{P}(\mathcal{T})$$

che associa a $\varepsilon, a, t \overset{\tau}{\mapsto} \tau(\varepsilon, a, t)$ un insieme di tratti, ed una funzione o, a valori reali nell'intervallo $[-1, 1]$, parzialmente definita su $\mathbb{E}$ e $\mathcal{T}$,

$$o : \mathbb{E} \times Ag \times \mathbb{R}^+ \times \mathcal{T} \to [-1, 1] \cap \mathbb{R}$$

dove

$$o(\varepsilon, a, t, \tau) = x \qquad \text{se } \tau \in \tau(\varepsilon, a, t)$$

Il valore di x esprime il grado di espressione (o di non-espressione) [7] di τ secondo a, nelle voci enciclopediche che riguardano ε.

Note. Siano fissati un agente a ed un istante di tempo t. Sia $\varepsilon \in \mathbb{E}$ un qualche argomento e sia $T \triangleq \tau(\varepsilon, a, t)$ l'insieme dei tratti in questione. Assumiamo $T \neq \emptyset$; allora si può definire l'immagine di T tramite o in $[-1, 1]$ come:[8]

$$o(T) = \{x \in [-1, 1] : \exists\, \tau \in T : x = o(\varepsilon, a, t, \tau)\}$$

ma è più utile, fissato un ordinamento di T convenzionale, elencare esplicitamente l'insieme di tratti, cioè scrivere $T = \{\tau_1, \ldots \tau_d\}$ per qualche $d = d(\varepsilon, a, t)$; allora, con un abuso di notazione, si può definire una nuova funzione $o : \mathbb{E} \times Ag \times \mathbb{R}^+ \to \mathbb{R}^d$ che associa $\varepsilon, a, t \mapsto o(\varepsilon, a, t) = (o_1, \ldots, o_d)$ ad un vettore, in modo tale che

$$o_i = o(\varepsilon, a, t, \tau_i) \quad i = 1, \ldots, d$$

A seconda se la dipendenza dal tempo risulti chiara dal contesto, userò anche le due notazioni vettoriali $\underline{o}^{\varepsilon,a} = \underline{o}^{\varepsilon,a}(t) = o(\varepsilon, a, t)$, per evidenziare il fatto che il risultato della funzione o è un vettore di numeri reali. Simili notazioni appariranno per denotare il vettore di tratti $\underline{\tau}^{\varepsilon,a} = \underline{\tau}^{\varepsilon,a}(t) = \tau(\varepsilon, a, t)$. Chiameremo *opinione* di un agente qualsiasi numero reale compreso nell'intervallo $[-1, 1]$, ogni qualvolta sia inteso in congiunzione con un tratto. Un vettore di opinioni, associato cioè ad un *insieme* di tratti, sarà detto *punto di vista* di un agente.

[7] τ_i è associato ad o_i nel senso che, se $o_i > 0$, esso rappresenta il grado di espressione di τ_i (secondo a e riguardo all'argomento ε) nell'enciclopedia; se $o_i < 0$, allora esso è il grado di non-espressione per tutti i τ_i' in contrasto con τ_i. La relazione "essere in contrasto con" è definita su $\mathcal{T} \times \mathcal{T}$ e dipende dalle assunzioni sperimentali che si fanno nella costruzione del modello. Infine, se $o_i = 0$ allora l'agente non ha alcuna opinione in merito a τ_i (e ai tratti contrastanti τ_i).

[8] In realtà, si tratta dell'immagine dell'insieme $\{(\varepsilon, a, t, \tau) : \tau \in T\}$.

Consideriamo due agenti a_1 e a_2 ed una pagina ε per cui $d(\varepsilon, a_1, t) > 0$ e $d(\varepsilon, a_2, t) > 0$. Vorremmo poter confrontare i punti di vista di a_1 ed a_2, per cui sorge spontaneo chiedersi in che spazio "vivano" i due vettori $\underline{o}^{\varepsilon,a_1}$ e $\underline{o}^{\varepsilon,a_2}$. In realtà, questa operazione presuppone che siano confrontabili i relativi insiemi di tratti, poiché non ha senso confrontare due gradi di espressione quando questi si riferiscono a due tratti diversi! Detto ciò, è abbastanza ovvio che il numero di dimensioni che servono per caratterizzare uno spazio comune ad a_1 e a_2 equivalga al numero di tratti *distinti* dei due punti di vista. Vogliamo dunque definire:

$$
\begin{aligned}
d(\varepsilon, \{a_i\}, t) &\triangleq d(\varepsilon, a_i, t) = |\tau(\varepsilon, a_i, t)| \quad i = 1, 2 \\
d(\varepsilon, \{a_1, a_2\}, t) &\triangleq |\tau(\varepsilon, a_1, t)| + |\tau(\varepsilon, a_2, t)| - |\tau(\varepsilon, a_1, t) \cap \tau(\varepsilon, a_2, t)|
\end{aligned}
$$

Cosicché, per ogni $\varepsilon \in \mathbb{E}$, usando la formula di inclusione-esclusione $d(\varepsilon, t) \triangleq d(\varepsilon, Ag, t)$ è ben definito.

Definition (Punti di vista accettabili)**.** Fissato un istante di tempo t, se $\underline{o}^{\varepsilon,a}$ è definito, allora esso appartiene con membership piena ad un insieme fuzzy A_a^ε su $\mathbb{R}^{d(\varepsilon,a,t)} \subseteq \mathbb{R}^{d(\varepsilon,t)}$, di tratti accettabili dall'agente a,

$$\forall a \in Ag,\ \exists\, \varepsilon \in \mathbb{E}\ :\ \mu_{A_a^\varepsilon}(\underline{o}^{\varepsilon,a}) = 1$$

L'insieme D_a^ε rappresenta al contrario gli argomenti inaccettabili ovvero incompatibili con il punto di vista dell'agente a.

Note. In generale, $D_a^\varepsilon \neq \overline{A}_a^\varepsilon$.

Note. La funzione di appartenenza di A_a^ε dovrebbe essere almeno unimodale; si potrebbe impiegare una qualsiasi normale multivariata per modellizzare forme ellittiche e simili. Un'alternativa potrebbe essere quella di impiegare un sistema esperto fuzzy, ma sorgerebbe il problema di gestire un numero di regole che cresce esponenzialmente nel numero delle dimensioni $d(\varepsilon, a, t)$. Darò un'occhiata a qualche sistema esperto di tipo neuro-fuzzy come le memorie associative fuzzy (le FAM di Bart Kosko), per vedere se possono essere impiegate per questo problema.

Una voce enciclopedica è una sorta di contenitore di testo per diversi sottoargomenti. Tuttavia, persino la voce ideale, secondo gli obiettivi editoriali di Wikipedia, dovrebbe riflettere diversi tratti non sempre in accordo tra essi, e riconducibili ai sottoargomenti presenti. Nel gergo di Wikipedia, si dice che i tratti presenti in una voce ideale dovrebbero riflettere tutti e soli i punti di vista *rilevanti*, relativi all'argomento in questione.

Definition (Voce enciclopedica "ideale")**.** Fissato un istante di tempo $t \in \mathbb{R}^+$, ad ogni $\varepsilon \in \mathbb{E}$ è associato un certo numero di sotto-argomenti, ciascuno dei quali è caratterizzato da un proprio insieme di tratti [9]. Formalmente, per ogni t una funzione insiemistica τ è definita su $\mathbb{E} \times \mathbb{E}$:

$$\tau : \mathbb{E} \times \mathbb{E} \times \mathbb{R}^+ \to \mathscr{P}(\mathcal{T})$$

essa associa a $\varepsilon, \varepsilon', t \mapsto \tau(\varepsilon, \varepsilon', t)$ un insieme di tratti, non nullo se e solo se $(\varepsilon, \varepsilon') \in E(\mathbb{E})$. La presenza di un tratto nel testo[10] della voce è data da un numero reale

[9] Ossia, da un suo punto di vista *generale*, sintesi dei punti di vista di più utenti editori della voce.

[10] Ovviamente l'espressione di un tratto può anche essere dovuta alla presenza di una immagine, o di altro contenuto multimediale.

nell'intervallo $[0,1]$; è dunque definita un'altra funzione o, parziale su $\mathbb{E} \times \mathbb{E}$ e su $\mathcal{T}$:

$$e : \mathbb{E} \times \mathbb{E} \times \mathbb{R}^+ \times \mathcal{T} \to [0,1] \cap \mathbb{R}$$

dove

$$(\varepsilon, \varepsilon') \in E(\mathbb{E}) \iff e(\varepsilon, \varepsilon', t, \tau) = x_\tau \qquad \forall \tau \in \tau(\varepsilon, \varepsilon', t)$$

Note. Sia fissato un istante di tempo t ed un argomento enciclopedico ε. Sia ε' un qualsiasi argomento enciclopedico che è collegato ad ε da un arco entrante; sia $T_{\varepsilon'} \triangleq \tau(\varepsilon, \varepsilon', t)$ l'insieme dei tratti che caratterizzano ε' come sottoargomento di ε. Allora è possibile definire l'immagine di $T_{\varepsilon'}$ tramite la funzione e in $[0,1]$:

$$e(T_{\varepsilon'}) = \{x \in [0,1] : \exists\, \tau \in T_{\varepsilon'} \,.\, x = e(\varepsilon, \varepsilon', t, \tau)\}$$

Anche in questo caso, è più utile usare una notazione vettoriale; *in primis* si consideri l'insieme dei vicini $N_\varepsilon = \{\varepsilon' : (\varepsilon, \varepsilon') \in E(\mathbb{E})\}$ scritto in forma esplicita:

$$N_\varepsilon = \{\varepsilon_1, \ldots, \varepsilon_D\}$$

dove $D = |N_\varepsilon|$, ed *in secundis* il multi-insieme T di tutti i possibili tratti

$$T = \{T_1, \ldots, T_D\}$$

in cui ciascun T_i è composto da un certo numero di tratti

$$T_i = \{\tau_1^i, \ldots, \tau_{d_i}^i\}, \quad i = 1, \ldots, D$$

per alcuni $d_i = d(\varepsilon, \varepsilon_i, t)$. Finalmente, con il solito abuso di notazione, è possibile definire un'altra funzione $e : \mathbb{E} \times \mathbb{R}^+ \to \mathbb{R}^{d^*}$, dove $d^* = \sum d_i$, la quale restituisce un vettore frutto della concatenazione dei D vettori di tratti di ciascun sottoargomento:

$$(\varepsilon, t) \overset{e}{\mapsto} e(\varepsilon, t) = (e_1, \ldots, e_{d^*})$$

in modo tale che, per $i = 1, \ldots, D$, si ha $k_i = 1, \ldots, d_i$, ed è allora definito $j \triangleq k_i + \sum_{h<i} d_h$ quale indice della generica componente

$$e_j = \tau(\varepsilon, \varepsilon_i, t, \tau_{k_i}^i)$$

del vettore. Le notazioni per $\underline{\tau}^\varepsilon$ e $\underline{e}^\varepsilon$ dovrebbero a questo punto essere intuitivamente ovvie.

Note. Quali sono i tratti che contraddistinguono ε_i come sotto-argomento di ε? Assumerò che $0 < d(\varepsilon, \varepsilon_i, t) \leq d(\varepsilon_i, t)$ poiché, a mio avviso, è privo di senso pensare che un argomento enciclopedico sia connotabile mediante tratti che non sono espressione di alcun punto di vista nella comunità. Dunque:

$$d^*(\varepsilon, t) = \sum_{i=1}^{|D|} d(\varepsilon, \varepsilon_i, t) \leq \sum_{i=1}^{|N_\varepsilon|} d(\varepsilon_i, t)$$

Ma $N_\varepsilon \subseteq \mathbb{E}$ mentre $d(\varepsilon_i, t)$ è la cardinalità di un insieme di tratti. C'è dunque una sostanziale differenza tra le due quantità d e d^*: la prima è la cardinalità dell'unione di tanti insiemi di tratti quanti gli agenti del sistema (al più), mentre la seconda è la cardinalità di un sottoinsieme del prodotto cartesiano $\mathbb{E} \times \mathcal{T}$.

Note. La definizione si limita ad asserire che ogni pagina è caratterizzabile in maniera univoca con un vettore di tratti. Sarà compito di un'altra parte del modello[11] dire quale sia il metodo che permette di identificare effettivamente tale vettore, a partire dai diversi punti di vista degli agenti sull'argomento in questione.

Come confrontare due voci? L'approccio che intendo utilizzare si basa sulla seguente idea intuitiva: una voce enciclopedica è caratterizzata dall'insieme dei suoi sottoargomenti, in altre parole dall'insieme di voci enciclopediche con cui è in relazione semantica per mezzo del grafo $G_{\mathbb{E}}$.[12] Siano ε, ε' due argomenti enciclopedici. Per confrontare i vettori $\underline{e}^{\varepsilon}$ e $\underline{e}^{\varepsilon'}$ dobbiamo sapere la dimensione di uno spazio vettoriale "comune". Siano $N_\varepsilon, T_\varepsilon$ (e similmente per ε') definiti come in precedenza.[13] L'insieme

$$N_\varepsilon \times T_\varepsilon \triangleq \{(\varepsilon_i, \tau) \ : \ \tau \in T_i, \ i = 1 \dots D\}$$

è formato da tutte le coppie argomento-tratto presenti nella voce di ε, ed infatti, fissato t, $|N_\varepsilon \times T_\varepsilon| = d^*(\varepsilon, t)$. A questo punto,

$$d^*(\{\varepsilon, \varepsilon'\}, t) \triangleq |(N_\varepsilon \times T_\varepsilon) \cup (N_{\varepsilon'} \times T_{\varepsilon'})|$$

conta il numero di coppie distinte, formate da un sotto-argomento e da un tratto, che sono presenti nella due voci enciclopediche, al tempo t. Questo ci permette di definire

$$d^*(t) \triangleq d^*(\mathbb{E}, t) \leq |\mathbb{E} \times \mathcal{T}|$$

come la dimensione dello spazio vettoriale in cui confrontare i vettori $\underline{e}^{\varepsilon}(t)$ ed $\underline{e}^{\varepsilon'}(t)$, per ogni $\varepsilon \in \mathbb{E}$, limitatamente all'istante di tempo t. A questo punto, è possibile definire

$$\hat{d} = \sup_t d^*(t)$$

come la dimensione complessiva del sistema.

Le definizioni presentate finora tentano di modellizzare il significato di una voce. In qualità di documento ipertestuale, ogni articolo possiede delle ben precise caratteristiche morfologiche. Le seguenti definizioni provano dunque a modellizzare gli aspetti importanti di una pagina in qualità di *significante*.

Definition (Stile di un agente)**.** Fissato un certo numero d' di caratteristiche morfologiche per i documenti ipertestuali di cui è composta Wikipedia, è possibile vedere lo stile editoriale di un agente come un insieme di valori che descrivono il modo in cui costui o costei [14] apporta modifiche ai documenti. Per ogni agente a e per ogni argomento ε è definita una funzione vettoriale

$$\mu : Ag \times \mathbb{E} \times \mathbb{R}^+ \to \mathbb{R}^{d'}$$

[11] Più precisamente, per mezzo della definizione effettiva di o e di e. Al limite si può accettare un approccio puramente oggettivista e dire che queste funzioni non dipendono in alcun modo dai vari $\underline{\tau}^{\varepsilon,a}$ e $\underline{o}^{\varepsilon,a}$.

[12] È sempre possibile presentare un esempio con due voci i cui argomenti enciclopedici non hanno assolutamente nulla a che fare l'uno con l'altro, ma che possono essere in relazione con lo stesso insieme di sottoargomenti. Queste due pagine, applicando la mia nozione intuitiva di similarità, risulterebbero "simili" o "confrontabili" quando nella realtà non lo sono minimamente. Tuttavia, se $\mathbb{E}$ è abbastanza grande, dovrebbe essere sempre possibile discriminare aggiungendo un sottoargomento di pertinenza a ciascuna delle due voci che rende distinte le due voci enciclopediche. Ribaltando questa critica, una nozione di similarità come quella che sto proponendo sembra realistica: se la comunità non è in grado di distinguere tra due voci enciclopediche perché legate ad argomenti enciclopedici *sotto-specificati*, allora è verosimile che tali argomenti non rivestano grande importanza all'interno dell'enciclopedia.

[13] Ovviamente, N_ε e T_ε dipendono da t, ma lo ometto dalla notazione per non appesantire troppo quest'ultima.

[14] Al limite anche un programma BOT!

che associa $a, \varepsilon, t \overset{\mu}{\mapsto} \mu(a, \varepsilon, t)$ e useremo il simbolo di vettore $\underline{\mu}^{\varepsilon,a} = \underline{\mu}^{\varepsilon,a}(t) \in \mathbb{R}^{d'}$ per riferirci al risultato di μ quando prende i parametri a, ε e t.

In maniera simile a quanto fatto per una voce ideale, definiamo la forma ideale che una pagina ipertestuale di Wikipedia dovrebbe avere, in base agli obiettivi editoriali fissati dal progetto.

Definition (Forma "ideale" di una voce enciclopedica)**.** Le grandezze di stato di carattere strutturale di una pagina sono soggette ad una dinamica nel tempo, dovuta al processo editoriale della voce; la funzione μ associa, in ogni istante di tempo t, all'argomento ε un vettore di numeri reali

$$\mu : \mathbb{E} \times \mathbb{R}^{+} \to \mathbb{R}^{d'}$$

cioè $\varepsilon, t \overset{\mu}{\mapsto} \mu(\varepsilon, t)$, dove al solito $\underline{\mu}^{\varepsilon} = \underline{\mu}^{\varepsilon}(t) \in \mathbb{R}^{d'}$ indica il risultato di μ con parametri ε e t; useremo anche il simbolo $\underline{\mu}$ ogni qualvolta la dipendenza da ε stesso risulterà chiara dal contesto.

Note. Il numero di tratti $d^{*}(\varepsilon, t)$ di una pagina ε dipende dal tempo t poiché la comunità può accogliere utenti portatori di nuovi tratti; si possono anche perdere tratti se dalla comunità vengono a mancare tutti gli utenti che rendevano un punto di vista rilevante per ε. Detto questo, anche la morfologia delle pagine di Wikipedia è cambiata nel corso degli anni, ma questa è standardizzata per tutte le pagine e sicuramente può essere modellizzata come un processo con dinamica lenta, se non esogena. Per questo d' non dipende né dal tempo, né tantomeno da ε.

Ecco un esempio concreto per illustrare le definizioni:

Example. Sia $\varepsilon \equiv$ "GEORGE BUSH", ed i suoi sottoargomenti, ad un certo istante di tempo, siano:

- Biografia
- Guerra in Iraq
- Presidenza degli Stati Uniti d'America

Ciascuno dei quali può essere caratterizzato da un tratto positivo e da uno critico, per un totale di $d(\varepsilon) = 6$. La morfologia di ε può essere:

- numero di caratteri nel testo
- numero di fonti (i.e. link esterni al sistema)
- numero di immagini.

Con le precedenti definizioni ho voluto tentare di modellizzare il fatto che la linea editoriale di Wikipedia esiste, ed è dettata dalla stessa comunità di utenti del progetto. Le seguenti definizioni catturano invece il fatto che la pagina "reale" di qualunque voce enciclopedica sottostà ad un'evoluzione dovuta al processo editoriale collaborativo, e che non necessariamente la pagina di una voce può essere sempre in aderenza al suo standard qualitativo "ideale"; ció avviene sia dal punto di vista del contenuto, che della struttura morfologica della voce.

Wikipedia è, dal punto di vista "fisico", un insieme di documenti ipertestuali collegati tra loro. Il grafo del wiki è il nome che viene comunemente dato a questo struttura; tale grafo viene creato in maniera collaborativa durante il processo di creazione di una pagina wiki.

Definition (Grafo del wiki)**.** $G^{\mathsf{w}}_{\mathbb{E}} \triangleq (V^{\mathsf{w}}_t(\mathbb{E}), E^{\mathsf{w}}_t(\mathbb{E}))$, dove $V^{\mathsf{w}}_t \subseteq \mathbb{E}$, $E^{\mathsf{w}}_t \subseteq V^{\mathsf{w}}_t \times V^{\mathsf{w}}_t$ è il grafo dei collegamenti ipertestuali tra pagine wiki.[15]

Essendo basata sul software Mediawiki, gli utenti di Wikipedia hanno accesso a tutte le versioni precedenti di ciascun articolo, e possono riportare una voce ad una versione precedente per mezzo di una operazione chiamata *rollback*. La pagina wiki di un argomento enciclopedico è quindi una successione di voci enciclopediche e di relative morfologie.

Definition (Versione)**.** Sia $\varepsilon \in \mathbb{E}$ un argomento enciclopedico; in maniera del tutto simile alle definizioni di voce enciclopedica "ideale" e di forma "ideale" di una voce enciclopedica sono definite le funzioni:

$$\begin{aligned} \tau^{\mathsf{w}} &: \mathbb{E} \times \mathbb{E} \times \mathbb{R}^+ \to \mathscr{P}(\mathcal{T}) \\ e^{\mathsf{w}} &: \mathbb{E} \times \mathbb{E} \times \mathbb{R}^+ \times \mathcal{T} \to [0,1] \cap \mathbb{R} \\ \mu^{\mathsf{w}} &: \mathbb{E} \times \mathbb{R}^+ \to \mathbb{R}^{d'} \end{aligned}$$

Cosìcché i vettori $\underline{\tau}^{\varepsilon}_{\mathsf{w}}$, $\underline{e}^{\varepsilon}_{\mathsf{w}}$ e $\underline{\mu}^{\varepsilon}_{\mathsf{w}}$ hanno il loro significato intuitivamente ovvio. Definiamo una funzione w che associa ad una argomento ε e ad un istante di tempo t la tripla:

$$\varepsilon, t \mapsto w(\varepsilon, t) \triangleq (\underline{\tau}^{\varepsilon}_{\mathsf{w}}, \underline{e}^{\varepsilon}_{\mathsf{w}}, \underline{\mu}^{\varepsilon}_{\mathsf{w}})$$

Definition (Pagina wiki)**.** Una pagina di Wikipedia per l'argomento $\varepsilon \in \mathbb{E}$ al tempo t è la successione di tutte le sue versioni fino a t, $(w(\varepsilon, t) : t \geq 0)$ indicizzata da $t \in \mathbb{R}^+$.

Definition (Cronologia di una pagina)**.** Ogni modifica di una pagina può essere vista come il salto in un nuovo stato per cui, per ogni $t \in \mathbb{R}^+$, è presente un mapping $\varepsilon, t \mapsto n(\varepsilon, t) \in \mathbb{Z}^+$, cioè si possono individuare $n(\varepsilon, t)$ versioni *distinte* della pagina, o alternativamente, esistono n intervalli disgiunti

$$[0, t_1), [t_1, t_2), \ldots, [t_{n-1}, t_n), [t_n, t) \subseteq \mathbb{R}^+$$

dove $n = n(\varepsilon, t) + 1$, per ciascuno dei quali vale $t' \in [t_i, t_{i+1}) \Rightarrow w(\varepsilon, t') = w(\varepsilon, t_i)$. Al tempo t la cronologia delle versioni precedenti è dunque

$$V^{\varepsilon}(t) = \{w(\varepsilon, t_1), \ldots, w(\varepsilon, t_n)\}$$

Definition (Watchlist di una pagina)**.** Al tempo $t \in \mathbb{R}$, la watchlist sulla pagina wiki relativa all'argomento $\varepsilon \in \mathbb{E}$ è un insieme di agenti registrati a ricevere notifica di ciascuna modifica effettuata su $w(\varepsilon, t)$, cioè $W^{\varepsilon}(t) \subseteq Ag$.

Definition (Punto di vista rilevante)**.** *(da definire)* Intuitivamente dovrebbe corrispondere ad un cluster di vettori di punti di vista nello spazio euclideo a $\mathbb{R}^{d^*(\varepsilon,t)}$ dimensioni.

Definition. Elenco delle interazioni del sistema:

- Creazione di una pagina wiki
- Modifica di una pagina wiki
- (Proposta di) fusione di due o più pagine wiki
- (Proposta di) smembramento di una pagina wiki

[15] La w indica che si tratta di una pagina wiki reale

- Ripristino di una pagina ad una versione precedente (rollback)
- Creazione di un link ipertestuale nel grafo del wiki
- Eliminazione di un link ipertestuale nel grafo del wiki
- Modifica di una opinione di un agente
- Creazione di un punto di vista rilevante su una pagina
- Fusione di due o più punti di vista rilevanti su una pagina
- Eliminazione di due o più punti di vista rilevanti su una pagina
- Richiesta[16] di cancellazione veloce di una pagina
- Proposta di cancellazione di una pagina
- Richiesta di blocco di un utente
- Amministratori e azioni loro permesse. *(da aggiungere in un modello avanzato)*

[16]Una richiesta differisce da una proposta poiché richiede l'intervento di un amministratore per effettivamente eseguire la cancellazione, nel caso in cui si raggiunga il consenso in questo senso.