

Pojęcie informacji, ilość informacji, maszyny informacyjne (skrót referatu)

Marian Piotr Jankowski

Seminarium prof. dr. hab. Jerzego Pelca, IF UW
22 lutego 2008

A. Informacja

1. Na wstępie zdefiniujemy pojęcie *Systemu* – jako ustalonego zbioru Elementów (tworzących jedno lub więcej Uniwersów), z określonymi na nich wartościowanymi Atrybutami (Działania i Funkcje) – tworzącymi Stan Systemu, który spełnia pewne Relacje. Np.: • dwa sprzężone koła zębate; • integrator maszyny analogowej; • kotka Lola i kanapa, fotel, miska.
2. Definiujemy pojęcie: *Hipostaza Konwencjonalna* – jako zabieg metodologiczny umownego uznania za istniejącą (jakby realnie) fikcyjnej substancji (itp. bytu), będącej zastępczym nośnikiem danej Własności Obiektów pewnej klasy. Przy tym metaforyczny rodzaj tej substancji (np. ciecz, gaz, kryształ, struna, ..) powinien zachowywać cechy strukturalne (modelować) tej Własności.
3. **Ogólna definicja pojęcia: Informacja** – to Własność pewnych składowych Stanu Świata (jako Systemu, jego fragmentu), wartościujących pewne Zdanie Pytające, polegająca na możliwości wartościowania na tej podstawie innego Zdania Pytającego, dotyczącego innych składowych Stanu Świata (innego fragmentu), wynikająca ze spełniania przez Stany Świata Relacji wiążących wartości tych Stanów.
4. Definicja ta pośrednio postuluje obecność *Podmiotu Testowego*, formułującego i wartościującego Pytania. Nie jest to jednak wymóg realnej jego obecności, a jedynie badania konsekwencji potencjalnej jego obecności i aktywności.
5. Wykazane cechy strukturalne tak zdefiniowanej Informacji jako Własności, pozwalają zinterpretować ją w trybie Hipostazy Konwencjonalnej jako Substancję – ciecz idealną, przepływającą ze stratą od Systemu Źródłowego, kolejno przez łańcuch następujących Systemów. Uprawnia to do mówienia o „przepływie” (itp.) Informacji.
6. *Dopuszczalne Odpowiedzi* na Pytanie to pewne zbiory Stanów Systemu – zdarzenia losowe. Punktowe wartości tych Stanów – wyniki (zdarzenia) elementarne – będziemy nazywać *Danymi*. Najprostsza, ogólna postać takiej Odpowiedzi – zdarzenia losowego dla składowych Stanu – liczb rzeczywistych, to ujemna półoś liczbowa zakończona ustalonym punktem (jak w dystrybucji).

7. Uogólnieniem pojęcia deterministycznej Odpowiedzi na Pytanie jest pojęcie Zwiększenia Wiedzy Istotnej Podmiotu Testowego, czyli zmniejszenie jego *Entropii Istotnej*. Istotnej – czyli odniesionej do *Wymaganej Dokładności*, będącej częścią Pytania.
8. Stąd, dla łańcucha Systemów, **uszczegółowiona definicja pojęcia: Informacja** – to Własność Stanu każdego z ciągu powiązanych (sprzężonych) Systemów, która rozchodząc się wzdłuż ciągu w formie przepływu stratnego, pozwala na podstawie znajomości Stanu pewnego Systemu ciągu zmniejszyć Entropię Istotną Zbioru Odpowiedzi Dopuszczalnych ustalonego Zdania Pytającego, dotyczącego Stanu Systemu Początkowego tego ciągu.
9. Zatem, **uproszczona, syntetyczna charakteryzacja pojęcia: Informacja** – to zdolność dostarczenia Wiedzy o Odpowiedzi na Pytanie dotyczące pewnego Obiektu, zawarta w innym Obiekcie.

B. Entropia

1. Dla przejrzystości wywodu, bez większej utraty ogólności, dalej Stany uznajemy za jednowymiarowe. Przyjmując jako dane:
 - bezwarunkową gęstość prawdopodobieństwa Stanu Systemu początkowego: $\varphi_0(x_0)$;
 - warunkowe gęstości prawdopodobieństw Stanów Systemów kolejnych: $\varphi_{n|n-1}(x_n|x_{n-1})$, $n = 1, 2, \dots$;
 przyjmując jako bezpośrednio niezależne Stany Systemów niesąsiednich (odległych), po kolei prezentowane są formuły na:
 - przenoszenie gęstości prawdopodobieństw wzdłuż ciągu Systemów: $\varphi_{n|0}(x_n|x_0)$, $n = 2, 3, \dots$;
 - odwrócenie tak otrzymanej gęstości wg tw. Bayesa: $\varphi_{0|n}(x_0|x_n)$, $n = 1, 2, 3, \dots$
2. Wprowadzamy *mechanizm filtru liniowego* z dwustronną *funkcją filtrującą* $g(v)$, tworzącą spłot z gęstościami prawdopodobieństw $\varphi(x)$ pod logarytmami wzorów na entropię. Gęstości po filtracji oznaczane będą ogólnie: $^*\varphi(x)$. Funkcja filtrująca $g(v)$ musi spełniać warunki:
 - być kawałkami ciągła;
 - być nieujemna;
 - być unimodalna w $v = 0$;
 - być unormowana: $g(0) = 1$;
 - mieć ograniczone nośniki dla każdego dodatniego poziomu;
 - być skończenie całkowna – o wartości całki δ .

Np.:

- (a) funkcja dzwonowa Gaussa;
 - (b) funkcja prostokątna.
3. Filtr liniowy poprzez „szerokość” funkcji δ wprowadza do Teorii Entropii i Informacji pojęcie Wymaganej Dokładności – konieczne dla zmiennych ciągłych i bardzo potrzebne dla zagęszczonych dyskretnych. Zalety w porównaniu z klasyczną definicją – dalej. Entropię zdefiniowaną tradycyjnie dla rozkładów ciągłych nazwiemy *Pseudoentropią*.

4. Definiujemy *Entropię Istotną* ${}^*H_0(X_0)$ zmiennej losowej Stanu X_0 (ciągłej lub dyskretnej – reprezentowanej dystrybucjami Diraca),
5. Wprowadzona definicja czyni Entropię Istotną:
 - (a) nieujemną,
 - (b) zerową dla rozkładów jednopunktowych,
 - (c) niemianowaną (bezwymiarową),
 - (d) niezmienniczą względem skali,
 - (e) niezmienniczą względem przekształceń jednoznacznych, ciągłych,
 - (f) nieskończoną w granicy z Wymaganą Dokładnością: $\delta \rightarrow 0$ (\rightarrow nieskończona rozdzielczość),
 - (g) zerową w granicy z Wymaganą Dokładnością: $\delta \rightarrow \infty$ (\rightarrow rozkład jednopunktowy).

Uzyskuje w ten sposób wszystkie podstawowe cechy klasycznej Entropii zdefiniowanej dla rozkładów dyskretnych i uznawane dla Entropii za konstytutywne, a **nie posiadane przez klasyczną Entropię zdefiniowaną dla rozkładów ciągłych** – Pseudentropię, mimo zwodniczej analogiczności odpowiadających wzorów. (A może właśnie dlatego!).

6. Entropia Istotna j.w. oferuje **formułę definicyjną ogólną dla rozkładów ciągłych i dyskretnych**. Tradycyjna Entropia dla rozkładów dyskretnych jest jej granicą dla prostokątnego $g(v)$ i $\delta \rightarrow 0$. i odwrotnie, Entropia Istotna jest granicą dla zagęszczającej się klasycznej Entropii dyskretnej.
7. Wprowadzamy *Punktową, Warunkową Entropię Istotną* a priori: ${}^*H_{0|n}(X_0|X_n)$, dla losowego X_n . Jest to zmienna losowa.
8. Wprowadzamy *Przeciętną, Warunkową Entropię Istotną* a priori: ${}^*H_{0||n}(X_0||X_n)$, jako wartość oczekiwaną Entropii Punktowej j.w. po X_n . Nietypowe oznaczenie „||” zastosowano, gdyż nie jest to warunkowa wartość oczekiwana w sensie ścisłym.
9. Entropie Istotne, będąc miarą nieokreśloności względem Wymaganej Dokładności, wyznaczają stopień Niewiedzy Istotnej – niezaniedbywalnej (niepomijalnej) z punktu widzenia kryterium dokładności Pytania. Wszelkie nadmiary tak ustalonej precyzji podana definicja wyklucza z rachunku. Jest to bardzo cenna jej cecha – odzwierciedlająca wymogi praktyki.

C. Ilość Informacji

1. Definiujemy *Ilość Informacji Istotnej a priori* (przed doświadczeniem): ${}^*I(X_0, X_n)$, jaką o Stanie X_0 może dostarczyć potencjalna znajomość Stanu X_n – jako odpowiednie zmniejszenie stopnia Niewiedzy Istotnej, czyli różnicę Entropii:

$${}^*I(X_0, X_n) = {}^*H_0(X_0) - {}^*H_{0|n}(X_0|X_n).$$

2. *Ilość Informacji Istotnej a posteriori* (po doświadczeniu), jaką o Stanie X_0 dostarcza faktyczna znajomość Stanu X_n – uzyskamy zamieniając w powyższym wzorze ${}^*H_{0|n}(X_0|X_n)$ na ${}^*H_{0|n}(X_0|X_n)$.

3. Za *Całkowitą Ilość Informacji* – $I(X_0, X_n)$ przyjmujemy maksymalne zmniejszenie Entropii, przy wszelkich możliwych Dokładnościach Wymaganych, tj.: $I(X_0, X_n) \geq *I(X_0, X_n)$. Można udowodnić, że maksimum to osiągnięte jest dla $\delta = 0$. Wówczas jest ono równe Ilości Informacji wyznaczonej z klasycznego wzoru, zarówno dla dyskretnych, jak i dla ciągłych rozkładów.
4. Przy bezpośredniej możliwości obserwacji X_0 , tj. $X_n = X_0$ zachodzi: $*H_{0|0}(X_0|X_0) = 0$. Natomiast $I(X_0, X_0) = \infty$. a więc, dowolny System posiada sam o sobie nieskończoną Ilość Informacji.

D. Maszyny Informacyjne

1. Realne Systemy dynamiczne faktycznie przetwarzają Dane – czyli swoje Stany. Czyniąc to, formują jednak swoją zdolność do określania Odpowiedzi na Pytanie – dla Podmiotu Testowego, czyli przetwarzają Informację. Dlatego Systemy te nazywać będziemy *Maszynami Informacyjnymi*.
2. Szczególną, niezwykle ważną klasą Maszyn Informacyjnych są *Maszyny Krujące Informację*. To Maszyny Informacyjne, które pozwalają uzyskać (tworzyć) Wiedzę – Odpowiedzi na Pytania, dla których Dane estymujące są bezpośrednio całkowicie niedostępne.
3. Maszyny Krujące Informację są Systemami dynamicznymi, w których strukturze spełnione są Relacje, będące odzwierciedleniem (modelem) Wiedzy o Systemie badanym – tj. Relacji w nim spełnianych.
4. Np.:
 - Odtworzenie głębi sceny w widzeniu obuocznym (zwierzęta);
 - Odtworzenie poglądów rozmówcy na podstawie wrywkowych jego wypowiedzi (ludzie);
 - Odtworzenie ewolucji struktur rozmieszczenia galaktyk (MC);
 - Prognoza/Odtworzenie wysokości lotu rakiety Space Ship One (MA).
5. Maszyny Informacyjne można ogólnie podzielić na Informacyjne Maszyny Proste i Złożone. Maszyny Proste oparte są na Modelu Prostem – tj. takim, w którym Punktowi, Funkcji, Działaniu, Relacji w Algebrze Systemu Oryginalnego odpowiednio odpowiadają Punkt, Funkcja, Działanie, Relacja w Algebrze Systemu Modelu. To np. maszyna analogowa, suwak logarytmiczny. Maszyny i Modele Złożone – to te, które nie są Proste. np. liczydło, umysł.