МЕТРИКИ СМЫСЛОВЫХ ПОДПРОСТРАНСТВ

Ивашенко В. П.
Кафедра интеллектуальных информационных технологий,
Белорусский государственный университет информатики и радиоэлектороники
Минск, Республика Беларусь
E-mail: ivashenko@bsuir.by

Рассматриваются подходы к построению метрик и соответствующих подпространств для случаев: интерпретируемого пространства скрытых признаков, выделяемых методами машинного обучения, включая нейросетевые методы, информационно-операционного пространства, ориентированого на интеграцию моделей с операционной семантикой большего и малого шага, подпространств на экстенсиональных замыканиях онтологических структур и графах состояний моделей обработки знаний, включая сильносвязные.

Введение

Разнообразие видов знаний и онтологических структур с семантикой различного вида в интеллектуальных системах требуют рассмотрения соответствующих моделей смысловых подпространств и методов их интеграции.

Разные виды смысловых пространств обладают разными структурными (топологическими) свойствами в некоторых случаях достаточными для задания метрики [1] в таких пространствах.

Сформулируем задачи, связанные с построением метрик и соответствующих смысловых подпространств [2]: (1) разработать онтологию видов смысловых пространств, которые позволяют соотносить известные смысловые пространства и подпространства общего смыслового пространства с их видами; (2) разработать модели метрических и топологических смысловых пространств конечных структур с операционной семантикой и конечных сильносвязных онтологических структур, согласованных с понятием экстенсионала и экстенсионального замыкания; (3) сформулировать требования к моделям метрического конечного операционно информационного смыслового пространства и метрического смыслового пространства протомультиверсальных чисел, согласующихся с расстоянием на действительных числах.

Под согласованностью топологического и метрического пространства понимается: каждая окрестность топологического пространства является объединением либо – конечным пересечением шаров, радиусы которых выражены метрикой относительно их центров.

I. Смысловые пространства и их виды

Онтология видов смысловых пространств (СП) представлена на рисунке 1.

Важным является вопрос, связанный с выражением понятия бесконечности в СП. Понятие бесконечности на некотором конечном этапе не может быть выражено конструктивно (актуально), поэтому бесконечность может быть сведена к недетерминизму. Однако при технической реализации, даже в случае недетерминизма (между конечным и бесконечным), в памяти интеллектуальной системы может быть зафиксирована не более чем счётная бесконечность. Несчётная бесконечность может быть выражена как недетерминизм возможного. На логическом уровне это может быть представлено средствами модальных логик: $\Diamond G \Diamond \bigcirc \top$; $\Diamond G \left(\Diamond \bigcirc \alpha \wedge \Diamond \bigcirc (\neg\alpha)\right)$.

При этом, данным выражениям могут соответствовать конечные циклические структуры.

II. Метрические подпространства и их метрики

Метрика и топология онтологических структур [2], включая сильносвязные. Пусть дано топологическое пространство $\langle X,S \rangle$. Обозначим топологическое замыкание множества M как $[M]_S$. Топологически сильносвязным множеством (классом неотделимости) будем называть класс отношения эквивалентности $[\{y\}]_S = [\{x\}]_S$.

Определим $C \in \left(2^X\right)^X$: $C(x) = \{y \,|[\{x\}]_S = [\{y\}]_S \}$. Допустим, что на классе $C(x)$ определена мера $\pi \in V^{C(x)}$, тогда введём меру $\mu \in V^X$ такую, что:

$$\mu(x) = 2*(\alpha-1)*\lim_{z\to\sup\{\pi(y)|y\in C(x)\}} \frac{\pi(x)}{z} + \\ +2*\sup\left(\{\mu(y)\,|[C(y)]_S \subset [C(x)]_S\}\right),$$

где $\alpha \in V$ и $\alpha > 1$, а $V \subseteq \mathbb{R}$.

При этом пусть $\pi(x) = 0$ при $\{x\} = C(x)$.

Определим метрику рекурсивно следующим образом:

$$\upsilon(\langle x,z\rangle) = \\ = \begin{cases} |\mu(x)-\mu(z)|\,|\eta(\langle x,z\rangle) \\ \inf\left(\{\upsilon(\langle x,y\rangle)+\upsilon(\langle y,z\rangle)\,|y\in X\}\right)|\neg\eta(\langle x,z\rangle) \end{cases}$$
$$\eta(\langle x,z\rangle) = ([\{z\}]_S \subset [\{x\}]_S) \vee ([\{x\}]_S \subset [\{z\}]_S)$$

Протомультиверсальные числа [2]. Для модели протомультиверсальных чисел $\langle N,L,R,F,M,A\rangle$ введём признаки:

$$\tau(\langle\chi,\gamma\rangle) = \frac{\chi+\gamma}{2};\sigma(\langle\alpha,\beta\rangle) = |\{\emptyset\,|\alpha\}|$$

и меры:

$$\delta(\langle\chi,\gamma\rangle) = \chi-\gamma;\lambda(\langle\alpha,\beta\rangle) = \sigma(\langle\alpha,\beta\rangle) = |\{\emptyset\,|\alpha \veebar \beta\}|$$

Выразим метрику на мультиверсальных числах:

$$\rho\left(\langle m,n\rangle\right)=\psi_1\left(\langle m.top,m.bottom,n.top,n.bottom\rangle\right)\odot$$
$$\odot\psi_2\left(\langle m.up,m.down,n.up,n.down\rangle\right)\odot$$
$$\odot\psi_3\left(\langle m.hi,m.low,n.hi,n.low\rangle\right)\odot\psi_4\left(\langle m,n,A\rangle\right)$$

$$\psi_1\left(\langle\alpha,\beta,\chi,\gamma\rangle\right)=$$
$$=\psi_6\left(\langle\varepsilon\left(\{\alpha\}\right),\varepsilon\left(\{\beta\}\right),\varepsilon\left(\{\chi\}\right),\varepsilon\left(\{\gamma\}\right)\rangle\right)$$
$$\psi_2\left(\langle\alpha,\beta,\chi,\gamma\rangle\right)=\odot_{f\in F}$$
$$\odot_{r\in R}\,\psi_1\left(\langle\alpha\left(f\right)\left(r\right),\beta\left(f\right)\left(r\right),\chi\left(f\right)\left(r\right),\gamma\left(f\right)\left(r\right)\rangle\right)$$
$$\psi_6\left(\langle\alpha,\beta,\chi,\gamma\rangle\right)=$$
$$=\varphi_1\left(\langle\delta,\alpha_2,\beta_2,\chi_2,\gamma_2\rangle\right)\odot\varphi_2\left(\langle\tau,\alpha_2,\beta_2,\chi_2,\gamma_2\rangle\right)\odot$$
$$\odot\varphi_3\left(\langle\lambda,\alpha_1,\beta_1,\chi_1,\gamma_1\rangle\right)\odot\varphi_4\left(\langle\sigma,\alpha_1,\beta_1,\chi_1,\gamma_1\rangle\right)$$

$$\varphi_i\left(\langle\phi,\alpha,\beta,\chi,\gamma\rangle\right)=\left|\phi\left(\langle\alpha,\beta\rangle\right)-\phi\left(\langle\chi,\gamma\rangle\right)\right|$$

$$\alpha\odot\beta=\lim_{t\to p}\left(\sqrt[t]{\alpha^t+\beta^t}\right);\varepsilon\left(S\right)=\left\{\begin{array}{l}0\left|\emptyset=S\right.\\x\left|x\in S\right.\end{array}\right.$$

Функции ψ_3, ψ_4, ψ_5 определяются аналогично ψ_2.

Операционно-информационное пространство [2] задаётся отношением перехода R между конфигурациями из множества C, множеством операций O, множеством операторов K, множеством параметров P и множеством значений V. Рассмотрим переход между конфигурациями в информационно операционном пространстве. Каждый переход связан с выполнением одной или нескольких команд соответствующих операций. Каждая операция тратит время на чтение и изъятие входных данных, а также на обработку и запись выходных данных (значений параметров). Обозначим время работы операции o, исполняющей оператор k, над входным параметром q с входными данными i с выходным параметром p и выходными данными u через $t\left(\langle k,o,q,i,p,u\rangle\right)$. Время работы операции o зависит не менее чем линейно $c>1$ от соответствующего редакционного расстояния [1,2] $\rho\left(\langle i,u\rangle\right)$: $t\left(\langle k,o,q,i,p,u\rangle\right)\geq c*\rho\left(\langle i,u\rangle\right)$. Расстояние между конфигурациями c и s не менее максимального из времён сработавших операторов k_j в c: $\rho\left(\langle c,s\rangle\right)\geq\rho_1\left(\langle c,s\rangle\right)\ \geq\ \max\left(\left\{t\left(\langle k_j,o_j,q_{g_j},i_{g_j},p_{h_j},u_{h_j}\rangle\right)\left|j\right.\right\}\right)$. Если конфигурации c и s отличаются и\или множеством операторов (K_c,K_s), операций (O_c,O_s), параметров (P_c,P_s), дуг (E_c,E_s), соединяющих операторы, параметры и операции, то расстояние $\rho\left(\langle c,s\rangle\right)$ и квазиметрика $\zeta\left(\langle c,s\rangle\right)$ между ними не меньше:

$$\rho\left(\langle c,s\rangle\right)\geq\zeta\left(\langle c,s\rangle\right)\geq\kappa_1*\rho_1\left(\langle c,s\rangle\right)\odot$$
$$\odot\omega_2\left(\langle\rho_2,\left(E_s/E_c\right),\left(E_c/E_s\right)\rangle\right)\odot$$
$$\odot\omega_3\left(\langle\rho_3,\left(P_s/P_c\right),\left(P_c/P_s\right)\rangle\right)\odot$$
$$\odot\omega_4\left(\langle\rho_4,\left(K_s/K_c\right),\left(K_c/K_s\right)\rangle\right)\odot$$
$$\odot\omega_5\left(\langle\rho_5,\left(O_s/O_c\right),\left(O_c/O_s\right)\rangle\right);$$
$$k_i\geq 0;\omega_i\left(\langle\rho,\alpha,\beta\rangle\right)=\kappa_i*$$
$$*\min(\{\rho\left(x\right)|(f\in\beta^{\alpha})\wedge\left(f^{-1}\in\alpha^{\beta}\right)\wedge$$
$$\left(x\in f\cup\{\emptyset\}\times\beta/Dom\left(f^{-1}\right)\cup\alpha/Dom\left(f\right)\times\{\emptyset\}\}\right)$$

Как вариант, расстояние ρ_5 между операциями может быть оценено ($\rho_5=\delta$):

$$\delta\left(\langle\varphi,\psi\rangle\right)=\odot\left(\{0\}\cup\left\{\theta\left(\rho\left(\langle\chi,\gamma\rangle\right)\right)\left|\langle\chi,\gamma\rangle\in\varphi-\psi\right.\right\}\right).$$

Заключение

Разработана онтология видов СП, включая метрические и топологические. Разработаны модели метрического и топологического СП конечных структур с операционной семантикой и конечных сильносвязных онтологических структур, согласованных с понятием экстенсионала и экстенсионального замыкания. Сформулированы требования к моделям метрического конечного операционно- информационного СП и метрического СП протомультиверсальных чисел, согласующихся с расстоянием на действительных числах. Открытым остаётся вопрос об уточнении топологии и метрики временных онтологических структур.

1. Петровский, А. Б. Пространства измеримых множеств мультимножеств. М. : Поли Принт Сервис. – 2016. – 324 с.
2. Ivashenko, V. Approaches to the Study of Semantic Space and Integrated Logical Inference Models Using Similarity, Difference and Other Measures = Подходы к исследованию семантического пространства и моделей интегрированного логического вывода с использованием мер сходства, различия и других мер / V. Ivashenko // Open Semantic Technologies for Intelligent Systems (OSTIS) : сборник научных трудов / Белорусский государственный университет информатики и радиоэлектроники ; редкол.: В. В. Голенков [и др.]. – Минск, 2025. – Вып. 9. – С. 97–112.

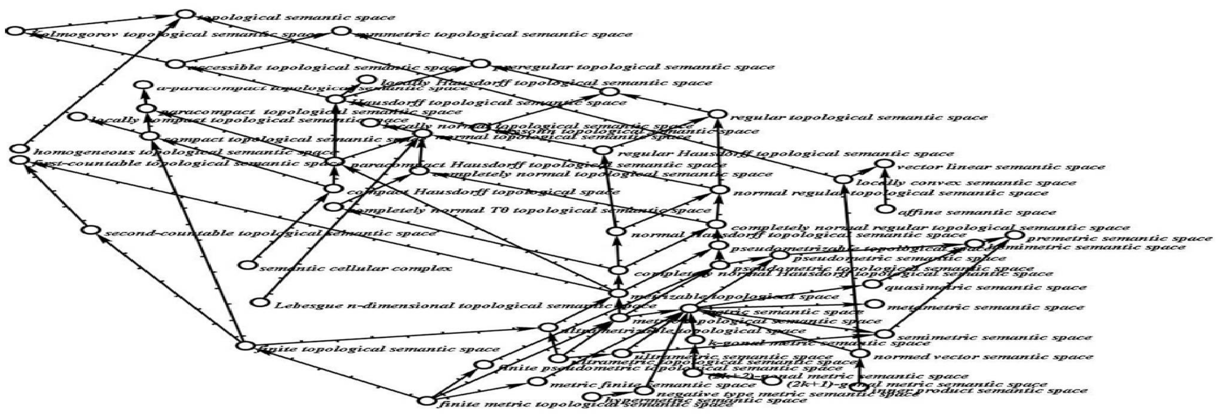

Рис. 1 – Онтология смысловых пространств