

Gráfelméleti eszközök az empirikus szociológia kumulatív felépítésének vizsgálatához

Bevezetés

A szociológiai jelenségek empirikus megragadása (a szociológiai vizsgálat) négy fő lépésben történik:

- I. A jelenséget leíró változók felvétele (hipotetikus modell).
- II. A jelenséget hordozó társadalmi objektumok kiválasztása (mintavétel).
- III. A felvett változók mérése a mintán.
- IV. A mérési adatok kiértékelése (empirikus modell).

E tárgykor részletes elemzésével számos irodalom foglalkozik (lásd pl. [11], [12]), így itt részletesebb kifejtésével nem foglalkozunk.

Ha azonban a szociológiai vizsgálatokat nem önálló egységeknek tekintjük, hanem egy társadalmi jelenség leírására szolgáló megismerési folyamat egy-egy fázisának, akkor többek között a következő alapvető problémák merülnek fel:

- a. A felvett változók halmaza két különböző vizsgálatban általában nem egyezik meg.
- b. Különböző vizsgálatok esetén (már az a. probléma következtében is) fellép a minták reprezentativitásának különbözősége, mint összehasonlítás ellen ható tényező.
- c. Különböző vizsgálatokban az azonosnak tekintett (azonos névvel ellátott) változók mérésében is sokszor eltérések mutatkoznak.
- d. A mérési adatok kiértékelésére a legkülönbözőbb módszereket alkalmazták.

A fenti problémák alapján, alapvető kérdésként adódik, hogy *mi biztosítja egy adott társadalmi jelenségre vonatkozó különböző vizsgálatok összehasonlíthatóságát, illetve hogyan vizsgálható és jellemezhető azok egymásraépülése (kumulativitása)?*

Az a., b., c., d. típusú problémákkal általában matematikai statisztikai megközelítésből foglalkoznak, így az eredmények is e területről származnak (lásd [5], [8], [9]), néhány kísérlettel eltekintve, amelyek strukturális modellezési eszközöket is javasolnak (lásd pl. [3], [7]).

Jelen dolgozatunkban a hangsúlyt olyan strukturális modell (gráf modell) ismertetésére helyezzük, amely az előzőekben feltett kérdésekre, illetve megválaszolásukra elvet és eszközt kínál.

1. A matematikai modell

Jelöljük egy adott J társadalmi jelenség leírására végzett szociológiai vizsgálatokat rendre a T_1, T_2, \dots, T_r jelekkel.

Ahhoz, hogy az egyes $T_i (i = 1, 2, \dots, r)$ vizsgálatok fogalmilag és műveletileg egyaránt matematikai modellbe ágyazhatók legyenek, definiálunk néhány szükséges fogalmat, majd ezek segítségével megadjuk a szociológiai vizsgálat matematikai definícióját. Ezek után mód nyílik a szociológiai vizsgálatok kumulativitásának pontos értelmezésére és definiálására.

Jelölje H a J jelenséget hordozó összes társadalmi objektumok halmazát. Ha például a J jelenség a bűnözés, akkor a jelenséget hordozó társadalmi objektumok a bűnelkövetők és a károsultak, ha a J jelenség a vállalatok közötti kooperáció, akkor a hordozó objektumok a vállalatok, intézmények. Megjegyezzük, hogy a H halmaz elemei a szociológiai vizsgálatok jelentős részében egyének.

Jelölje továbbá V a J jelenséget leíró összes változók halmazát, amelyről lényeges megjegyezni, hogy nem feltétlenül véges halmaz.

A továbbiakban az elemi definíciók jelölésére a $d1, d2, \dots$ jelöléseket fogjuk használni.

d1.

Kódhalmaz alatt egy véges karakterkészlet felett értelmezett, véges hosszúságú karaktersorozatokból (kódközből) álló halmazt értünk.

1.1. Definíció

Egy $v \in V$ változó H halmazra vonatkozó mérésén, olyan egyértelmű m leképezést értünk, amely az alábbiak szerint definiálható:

$$(1.1) \quad m : v \times H \rightarrow K_v,$$

azaz

$$(1.2) \quad \forall h \in H \Rightarrow \exists! k \in K_v : m(v, h) = k,$$

ahol K_v a $v \in V$ változóhoz tartozó kódhalmaz.

Nem térünk itt ki a mérés fogalmának és elméletének további elemzésére, mert igen messzire vezetne és e tárgykör jelenlegi helyzetének kimerítő tárgyalása található [5]-ben. Azonban felhívjuk a figyelmet arra, hogy a mérés típusát (szintjeit) pontosan az alkalmazott K_v kódhalmaz tulajdonságai szabják meg, amelyek nem biztos, hogy megegyeznek a v változó tényleges tulajdonságaival.

Az esetek nagy többségében (egyész nézetek szerint mindig), a K_v halmazról feltételezik, hogy az a valós számhalmaz valamely részhalmaza, ami nagy kísértés lehet a „szám” és a „szám karakter” (számjegy) tulajdonságainak azonosítására.

A bevezetésben leírt I., II., III. lépéseket és a fentieket figyelembe véve, a J jelenség megismerésére végzett tetszőleges $T_i (i = 1, 2, \dots, r)$ szociológiai vizsgálat megtervezése, egy (V_i, H_i, M_i) halmaz-triáddal jellemezhető, ahol

V_i : a T_i vizsgálatban felvett *változók halmaza*, melyre teljesül

$$(1.3) \quad V_i \subseteq V \text{ és } |V_i| < \infty.$$

$A \subseteq$ jellel a tartalmazási, a \subset jellel a szigorú tartalmazási relációt jelöljük.

H_i : a T_i vizsgálatban felvett *mintaelemek halmaza*, melyre teljesül

$$(1.4) \quad H_i \subseteq H.$$

A H_i halmaz megkonstruálását nevezzük *mintavételnek*. A H_i és a H halmaz viszonya jellemzi a minta típusát, például $H_i = H$ esetén teljes körű vizsgálatról beszélünk. Egyébként bizonyos kijelölt (fő) változók szerinti reprezentatívítást szoktunk a H_i mintától megkövetelni, vagy a mintát véletlenszerűen választjuk H -ból.

M_i : a V_i halmazbeli változókhoz tartozó, H_i halmazra vonatkozó mérési eljárások halmaza (lásd 1.1. definíció), azaz

$$(1.5) \quad \forall v_{ij} \in V_i \stackrel{\text{def.}}{\implies} \exists ! m_{ij} \in M_i : \forall h \in H_i \implies \exists ! k \in K_{v_{ij}} : m_{ij}(v_{ij}, h) = k,$$

ahol $K_{v_{ij}}$ a $v_{ij} \in V_i$ változóhoz rendelt kódhalmaz.

A szociológiai vizsgálat elvégzése pontosan azt jelenti, hogy a kiválasztott mintán elvégezzük a kiválasztott változók szerinti mérési eljárásokat. Ily módon egy adathalmazt állítunk elő, amely hipotézisünk szerint a J jelenség (közelítő) leírására alkalmas. Az eddig bevezetett fogalmainkkal ez a következőként írható le.

Mivel

$$(1.6) \quad \begin{array}{ccccccc} v_{i1} \times H_i & \cup & \dots & \cup & v_{ij} \times H_i & \cup & \dots & \cup & v_{i|V_i|} \times H_i & = & V_i \times H_i, \\ m_{i1} \downarrow & & & & m_{ij} \downarrow & & & & m_{i|V_i|} \downarrow & & \\ K_{v_{i1}} & & \dots & & K_{v_{ij}} & & \dots & & K_{v_{i|V_i|}} & & \end{array}$$

így a J jelenségre vonatkozó $T_i = (V_i, H_i, M_i)$ ($i = 1, 2, \dots, r$) szociológiai vizsgálat elvégzése az alábbi τ_i egyértelmű leképezésként fogható fel

$$(1.7) \quad \tau_i : V_i \times H_i \rightarrow K = K_{v_{i1}} \cup K_{v_{i2}} \cup \dots \cup K_{v_{i|V_i|}}.$$

Ekkor a T_i vizsgálat elvégzése során $|V_i| \cdot |H_i|$ mérést végzünk, amelyek eredményeként ugyanennyi adatot (kódot) nyerünk.

E fejezet további részében a kumulativitás fogalmának definíciójához jutunk el, a fogalom tartalmának érzékeltetése és a szükséges segédfogalmak bevezetése után.

1.2. Definíció

A V_i változóhalmaz *struktúráján* egy R trichotom relációt értünk, melyre

$$(1.8) \quad R \subseteq V_i \times V_i$$

Az R reláció tartalmát tekintve, a változók közötti viszonyokat, összefüggéseket (pl. ok-okozat) jelöl.

A következőkben gráf alatt mindig irányított gráfot értünk, az irányítatlan esetet külön jelezzük.

A 1.2. definíció, valamint a relációk és gráfok közötti ismert megfeleltetés (lásd pl. [4], [10]) alapján, bármely (V_i, R) változó-rendszerhez egyértelműen rendelhető egy $\Gamma_i = (P_i, E_i)$ gráf, ahol P_i a gráf szögpont, E_i pedig az élhalmazát jelöli. (A rendszer fogalmának fenti értelmezését lásd [2]-ben.)

Mindegyik P_i -beli p_{ij} szögpont, a megfelelő V_i -beli v_{ij} változót reprezentálja, az E_i élhalmaz pedig a következőképpen áll elő

$$(1.9) \quad (p_{iq} p_{iz}) \in E_i \Leftrightarrow v_{iq} R v_{iz} \quad (p_{iq}, p_{iz} \in P_i).$$

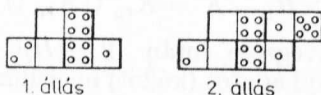
A fentiek alapján tehát a T_1, T_2, \dots, T_r vizsgálatokhoz hozzárendelhető egy $\Gamma_1, \Gamma_2, \dots, \Gamma_r$ gráf sorozat.

A kumulativitás fogalmával a vizsgálatok egymásraépülését szeretnénk kifejezni, ezért alapvetően két szempontra kell figyelemmel lennünk e fogalom definiálásánál:

- A kumuláló vizsgálatnak terjedelmében tartalmaznia kell a kumulált vizsgálatot. Ez azt jelenti, hogy csak olyan vizsgálat kumulálhat másikat, amely az adott jelenségből nagyobb részt ír le, mint a másik, azaz legalább eggyel több változó mentén vizsgálja az adott jelenséget.
- A kumuláló vizsgálatnak struktúrájában is tartalmaznia kell a kumulált vizsgálatot. Ez azt jelenti, hogy két különböző vizsgálatban csak akkor azonosak ténylegesen az azonos névvel ellátott változók, ha a közöttük levő viszonyok (R reláció) nem változnak meg csupán attól, hogy két különböző időpontban mértük őket.

A kumulativitás fogalmának intuitív megközelítésére tekintsünk két közismert játékot, a dominót és a sakkot.

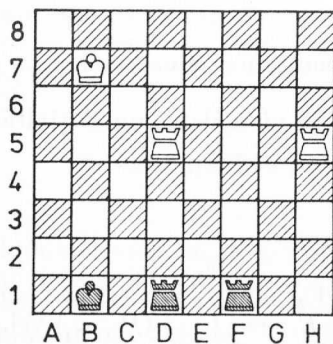
- A dominó játék esetében, bármely állásnál az újonnan letett dominó terjedelmében gyarapítja az előző állást oly módon, hogy a megelőző állás struktúráját (a dominók egymáshoz való viszonyait) az új struktúra magába foglalja (kumulálja). Tehát a dominó játékot nevezhetjük kumulatív játéknak (lásd 1. ábra).



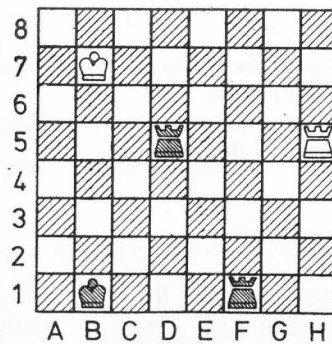
1. ábra

- A sakk játék esetében, bármely állásból a következő úgy jön létre, hogy az előző struktúrája (a figurák egymáshoz való viszonya) megváltozik. Két egymásutáni állásnál soha nem jöhet létre, hogy az utóbbi tartalmazná a teljes előző állást, hiszen ez feltételezné az előző állás változatlan-ságát, amiből következne, hogy nem jöhet létre az új állás. Esetleg egy adott állás után több lépéssel állhat elő olyan állás, amely az adott állást teljes egészében tartalmazza (lásd 2. ábra). Tehát a sakkot nem nevezhetjük kumulatív játéknak.

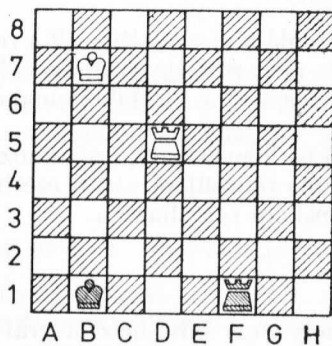
A következőkben használt gráfelméleti terminológiát az [1], [4]-ben használatos módon definiáljuk.



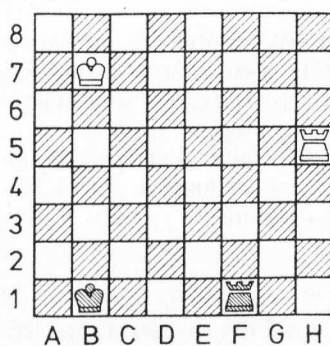
1. állás



2. állás



3. állás



4. állás

2. ábra

d2.

Ha egy G gráfból elhagyunk egy vagy több szögpontot mindazon élekkel együtt, amelyeknek egyik vagy mindkét végpontja az elhagyott szögpontok között van, akkor az így keletkező gráfot a G gráf részgráfiájának nevezzük.

d3.

Adott G gráfbeli pályának nevezzük a G -beli éleknek olyan egymásutánját (sorozatát), amelyben mindegyik él végpontja megegyezik az utána következő él kezdőpontjával és egyik szögpontot sem érintjük kétszer. Körpályának, vagy röviden körnek nevezünk egy olyan pályát, amelynek kezdő és végpontja megegyezik.

d4.

A G gráfbeli láncon értjük a G -beli élek tetszőleges egymáshoz kapcsolódó sorozatát, tekintet nélkül az élek irányítására.

d5.

A G gráf összefüggő, ha bármely két szögpontjához létezik a kettőt összekötő G -beli lánc.

Most megadjuk két tetszőleges szociológiai vizsgálat kumulativitásának matematikai definícióját.

1.3. Definíció

Jelöljük a T_i és T_j vizsgálatok, V_i illetve V_j változóhalmazainak struktúráit reprezentáló gráfokat rendre $\Gamma_i = (P_i, E_i)$ és $\Gamma_j = (P_j, E_j)$ -vel. A T_j vizsgálatot *kumulatív*nak mondjuk a T_i vizsgálatra nézve, ha

$$(1.10) \quad \Gamma_i \subset \Gamma_j, \text{ azaz } P_i \subset P_j \text{ és } E_i \subset E_j,$$

$$(1.11) \quad \Gamma_j \text{ összefüggő.}$$

Az (1.10) feltétel azt mondja ki, hogy a kumulált vizsgálat (T_i) rendelt gráf (Γ_i) , részgráfja a kumuláló vizsgálat (T_j) rendelt gráfnak (Γ_j) . Ez tehát a kumulativitás fogalmával szemben támasztott *a.*, *b.* követelményeket fogalmazza meg.

Az (1.11) feltétel szükségessége nem látható be közvetlenül, így ahhoz némi magyarázat szükséges. Az (1.10) feltétel ugyanis teljesülhet olyan esetben is, amikor például a T_j gráf a T_i gráfot komponensként tartalmazza.

d6.

Egy gráf olyan összefüggő részgráfját, amely nem bővíthető a gráf szögpontjainak, vagy éleinek hozzávételével úgy, hogy összefüggő maradjon, a gráf egy *komponensének* nevezzük.

Az említett esetben tulajdonképpen a T_j vizsgálat csak névlegesen egy vizsgálat, hiszen a különböző Γ_j -beli komponensekbe tartozó változók függetlenek a másik komponensbe tartozó változóktól.

Szemléletesen e helyzet úgy mutatható be (ismét a dominó játékot használva segédeszközként), mintha a dominó játékban, egy játszma belül, egy adott állásnál valaki az asztal egy másik részére helyezné el a következő dominót (azaz játszma közben új játszmat kezdene). Természetes, hogy ekkor nincs értelme a játszma folytatásáról beszélni, hiszen szélsőséges esetben, minden lépésnél tetszőleges helyre lehetne így dominót elhelyezni, ami magát a játékot szüntetné meg.

Visszatérve a szociológiai vizsgálatokhoz, a fenti T_j vizsgálat úgy fogható fel, mint egymástól független vizsgálatok egy időben való elvégzése, ahol a független vizsgálatok struktúrái a komponenseknek felelnek meg. Ekkor viszont nincs értelme egymásraépülésről (kumulativitásról) beszélni, ezért zárjuk ki ezt az esetet az (1.11) feltétellel.

Az 1.3. definíció alapján, adott J szociológiai jelenség esetén, a kumulativitás a T_1, T_2, \dots, T_r vizsgálatok halmazán értelmezett, bináris reláció. A továbbiakban jele: R_c .

A „ T_j vizsgálat kumulatív a T_i vizsgálatra nézve” reláció formalizált leírása: $T_j R_c T_i$.

A következőkben az R_c reláció tulajdonságaival, illetve ennek alapján az empirikus szociológiai megismerési folyamat struktúrájának elemzésével foglalkozunk.

2. Az empirikus megismerési folyamat struktúrájának elemzése

A kumulativitás (R_c reláció) lehetőséget ad két tetszőleges vizsgálat összehasonlítási problematikájának általánosítására, azaz egy J jelenség adott időpontig végzett teljes empirikus megismerési folyamatának elemzésére.

Tekintsük ugyanis az R_c reláció gráfját, azaz azt a $\Gamma_{R_c} = (P_{R_c}, E_{R_c})$ gráfot, melynek szögpontjai a T_1, T_2, \dots, T_r vizsgálatokat reprezentálják úgy, hogy a T_i vizsgálatnak a p_i szögpont felel meg, a gráf éleit pedig az alábbiak szerint definiáljuk:

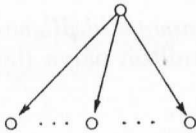
$$(2.1) \quad (p_i p_j) \in E_{R_c} \Leftrightarrow T_i R_c T_j.$$

A \subset reláció ismert tulajdonságai (irreflexív, antiszimmetrikus, tranzitív) alapján az 1.2. definícióból adódik, hogy az R_c reláció rendelkezik az irreflexív, antiszimmetrikus, tranzitív tulajdonságokkal. Ez azt jelenti, hogy a *kumulativitás* a vizsgálatok halmazán értelmezett *szigorú, rendezési reláció*.

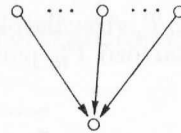
d7.

Egy tranzitív relációt reprezentáló gráfot *tranzitív gráfnak* nevezünk.

A tranzitív gráf körmentes, irányított gráf, melyben két szögpontot (p, q) akkor és csak akkor köt össze (pq) él, ha a gráfban van a p szögpontból a q szögpontba vezető pálya (ez a pálya lehet egy él is).



3. ábra



4. ábra



5. ábra



6. ábra

Lényeges tulajdonsága a tranzitív gráfoknak, hogy bármely részgráfjuk is rendelkezik a fenti tulajdonságokkal, azaz bármely részgráfjuk tranzitív gráf.

Az alábbi 3–6. ábrákon bemutatjuk azokat az elemi gráftípusokat, amelyekből minden fenti tulajdonságú Γ_{R_c} gráf felépíthető.

A 3. ábra elemi gráfjához tartozó megismerési struktúrát *összegző*, a 4. ábra

elemi gráfjához tartozót *kereső típusnak* nevezzük. Az 5. ábra megismerési struktúráját pedig *elemi ideális* struktúrának nevezzük.

Érdekes megfigyelní, hogy az elemi ideális típus, az összegező és kereső típusok közös határesetete.

d8.

Ha egy G gráf összes szögpontját megtartjuk, de elhagyjuk egy vagy több élét, akkor a G gráf egy *parciális gráfjához* jutunk.

d9.

Legyenek p_1, p_2, p_3 tetszőleges szögpontok egy adott $G = (P, E)$ tranzitív gráfban. Ha teljesül az alábbi három feltétel,

$$(2.2) \quad (p_1 p_2) \in E$$

$$(2.3) \quad (p_2 p_3) \in E$$

$$(2.4) \quad (p_1 p_3) \in E,$$

akkor a $(p_1 p_3)$ élt *tranzitív élnek* nevezzük.

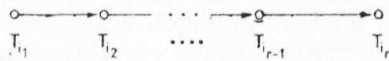
A továbbiakban jelöljük a Γ_{R_e} gráfból az *összes tranzitív élek elhagyásával* keletkező parciális gráfot Γ'_{R_e} -vel.

d9.

Egy n szögpontú gráf olyan pályáját, amely mind az n szögponton pontosan egyszer megy át, *Hamilton-pályának* nevezzük.

2.1. Definíció

A T_1, T_2, \dots, T_r vizsgálatokból álló *megismerési folyamatot ideálisnak* nevezzük, ha a hozzá tartozó Γ'_{R_e} parciális gráf egyetlen Hamilton-pálya (lásd 7. ábra).



7. ábra

d10.

Teljesnek nevezünk egy gráfot, ha bármely két szögpontját él köti össze.

1. Tétel

A T_1, T_2, \dots, T_r vizsgálatokból álló megismerési folyamat akkor és csak akkor ideális, ha a hozzá tartozó Γ'_{R_e} gráf tranzitív, teljes gráf.

Bizonyítás

A szükségesség bizonyításához azt kell belátnunk, hogy ha Γ_{R_e} tranzitív, teljes gráf, akkor annak létezik olyan Γ'_{R_e} parciális gráfja (amely az előzők-

ben leírtak szerint a tranzitív élek elhagyásával áll elő), amely egyetlen Hamilton-pálya.

Ehhez felhasználjuk Rédei László következő tételét ([6] II. fejezet 10. tétel):

„Egy legalább 2 szögpontú (irányítatlan) teljes gráf bármilyen irányítása révén nyert irányított gráfnak van irányított Hamilton-pályája.”

Rédei László tételének egyszerűbb megfogalmazása az alábbi:

Egy legalább 2 szögpontú irányított teljes gráfban mindig létezik Hamilton-pálya.

Folytatva tételünk bizonyítását, a Rédei tétel biztosítja legalább egy Γ_{R_c} -beli Hamilton-pálya létezését. Mivel Γ_{R_c} tranzitív gráf, így bármely olyan él, amely nem egy kiválasztott U Hamilton-pályához tartozik, tranzitív él. Ezeket a tranzitív éleket elhagyva tehát, a keletkező Γ'_{R_c} parciális gráf pontosan az U Hamilton-pálya lesz, azaz a megismerési folyamat ideális.

Az elegendőség bizonyításánál a megismerési folyamat ideális voltából indulunk ki.

Ekkor a 2.1 definíció szerint van Γ_{R_c} -beli Γ'_{R_c} Hamilton-pálya, amelyből az R_c reláció tranzitív tulajdonsága miatt következik, hogy Γ_{R_c} tranzitív és teljes. Ezzel a tételt beláttuk.

d11.

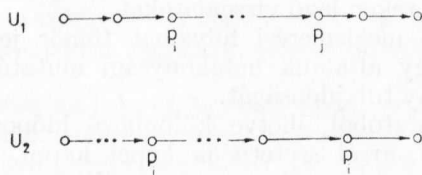
Legyen $G = (P, E)$ tetszőleges irányított gráf, és legyenek p és q tetszőleges P -beli szögpontok. Ha a G gráf bármely p, q szögpontpárja esetén teljesül, hogy azokat a (pq) és (qp) élek közül legfeljebb az egyik köti össze, akkor a G gráfot *antiszimmetrikusnak* nevezzük.

2. Tétel

Egy $G = (P, E)$ tranzitív, antiszimmetrikus, teljes gráf pontosan egy Hamilton-pályát tartalmaz.

Bizonyítás

Az idézett Rédei tétel biztosítja a G gráfban Hamilton-pálya létezését, így csak az unicitás bizonyítására szorítkozunk. Ehhez tegyük fel, hogy a G gráfban van két U_1 és U_2 Hamilton-pálya. Ekkor van legalább két olyan $p_i \in P$ és $p_j \in P$ szögpont, amelyek bejárási sorrendje U_1 -ben és U_2 -ben ellentétes (lásd 8. ábra).



8. ábra

Mivel G tranzitív gráf, így U_1 alapján $(p_i p_j) \in E$, U_2 alapján pedig $(p_j p_i) \in E$, ami G antiszimmetrikus volta miatt nem lehetséges.

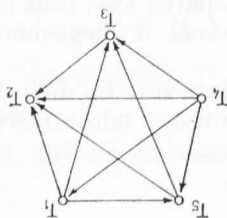
Ezzel a tételt beláttuk.

Az 1. és 2. tételekből adódik, hogy ideális megismerési folyamat pontosan akkor áll elő, ha a kumulativitás (R_c reláció) a T_1, T_2, \dots, T_r vizsgálatok halmazán trichotom, szigorú rendezési reláció.

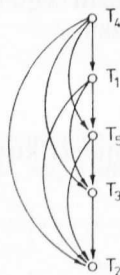
Egy ilyen ideális megismerési folyamathoz tartozó Γ_{R_c} gráfot mutat a 9. ábra.

Bármely Γ_{R_c} gráf, tranzitív gráf lévén rendelkezik azzal a tulajdonsággal, hogy szögpontjai szintekbe rendezhetők. Ez azt jelenti, hogy bármely szögpontjából csak a nála alacsonyabb szinten levő szögpontokba vezet él, a szögpontok megfelelő átrendezésével. (lásd [4]).

A 9. ábra gráfját szintekbe rendezve, a 10. ábra gráfját kapjuk, amely az előzővel izomorf.



9. ábra



10. ábra

Így világosabban látszik, hogy miért pont a 2.1. definíció szerinti megismerési folyamatot neveztük ideálisnak. Hiszen csak ekkor rendelkezik a folyamat azzal a tulajdonsággal, hogy bármely vizsgálat kumulálja (beépíti) az összes alacsonyabb szinteken levő vizsgálatokat.

A következőkben a megismerési folyamat tömör jellemzése (értékelése) céljából, megadunk egy általunk hatékonysági mutatónak nevezett összefüggést és ennek néhány tulajdonságát.

A hatékonysági mutatóból, illetve különböző időpontokban kiszámított értékeinek sorozatából, arról szeretnénk képet kapni, hogy a megismerési folyamat struktúrája mennyire konvergál az általunk ideálisnak tekintett megismerési struktúrához.

Jelöljük a J jelenség megismerésére a t_i ($i = 0, 1, 2, \dots$) időpontban rendelkezésre álló T_1, T_2, \dots, T_r vizsgálatok halmazán értelmezett R_c kumulatívítási reláció gráfját $\Gamma_{R_c}(t_i)$ -vel, ahol

$$(2.5) \quad \Gamma_{R_c}(t_i) = (P(t_i), E(t_i)).$$

Tekintsünk két t_i, t_j időpontot, ahol feltesszük, hogy $j > i \geq 0$.

Továbbá a $k(t_i, t_j)$ egész szám tegyen eleget az alábbi egyenlőtlenségnek:

$$(2.6) \quad k(t_i, t_j) \geq \begin{cases} 2, & \text{ha } i = 0 \\ 1, & \text{ha } i \geq 1. \end{cases}$$

Ennek segítségével rekurzív módon megadjuk bármely t_j időpontra az aktuálisan rendelkezésre álló vizsgálatok számát:

$$(2.7) \quad P(t_0) = 0.$$

$$(2.8) \quad \forall j > i \geq 1 \Rightarrow |P(t_j)| = |P(t_i)| + k(t_i, t_j).$$

Vezessük be a rövidebb írásmód kedvéért a $|P(t_i)| = n_i$ ($i = 0, 1, 2, \dots$) jelölést. Ekkor a J jelenség megismerési folyamatának t_i, t_j időintervallumbeli hatékonyságát a következő összefüggéssel adjuk meg.

$$(2.9) \quad H(t_i, t_j) = \frac{2(|E(t_j)| - |E(t_i)|)}{(n_i + k(t_i, t_j)) \cdot (n_i + k(t_i, t_j) - 1) - n_i(n_i - 1)}.$$

A $H(t_i, t_j)$ hatékonysági mutató konstrukciójánál két alapvető szempontot tartottunk szem előtt.

1. A megismerési folyamatot csak az adott t_i, t_j időintervallumban jellemezze. Azaz dinamikus legyen abban az értelemben, hogy a t_0, t_i időintervallumbeli hatékonyságrontó tényezők hatásától mentes legyen a $H(t_i, t_j)$ hatékonysági mutató.
2. A hatékonysági mutató ténylegesen (az adott időintervallumra) a megismerési folyamat adott szakaszának az ideálistól való eltérését jellemezze. Azaz maximális értékét az ideális, minimális értékét az ellenkező véglet esetén vegye fel.

Megmutatjuk, hogy a (2.9)-ben definiált hatékonysági mutató az előző 1., 2. követelményeket kielégíti.

d12.

Az olyan irányítatlan gráfot, amelyben bármely két szögpontot legfeljebb egy él köt össze, *egyszerű gráfnak* nevezzük.

Ismeretes, hogy egy n szögpontú egyszerű, teljes gráf éleinek száma $n(n-1)/2$ és ez egyben az n szögpontú egyszerű gráfok éleinek maximális számát is adja.

Figyelembe véve tehát az 1. tételt kimondhatjuk, hogy bármely t_i időpontban a $\Gamma_{R_c}(t_i)$ gráf éleinek száma, ideális megismerési struktúra esetén $n_i(n_i - 1)/2$, hiszen egy antiszimmetrikus, irányított teljes gráf (és $\Gamma_{R_c}(t_i)$ adott esetben ilyen) úgy áll elő, hogy az n_i szögpontú egyszerű, teljes gráf éleit megfelelő irányítással látjuk el (lásd pl. 9. ábra).

Ekkor (2.8) alapján a $t_j > t_i$ időpontbeli $\Gamma_{R_c}(t_j)$ gráf maximális élszáma pontosan

$$(2.10) \quad \frac{n_j(n_j - 1)}{2} = \frac{(n_i + k(t_i, t_j))(n_i + k(t_i, t_j) - 1)}{2}.$$

A $H(t_i, t_j)$ mutatóval szemben támasztott 1. követelményt figyelembe véve tekintsük úgy, hogy a t_i időpontbeli megismerési struktúra ideális, ekkor fennáll a következő tétel.

3. Tétel

Ha a t_0, t_i időintervallumban a megismerési folyamat ideális, azaz a $\Gamma_{R_c}(t_i)$ gráf irányított, antiszimmetrikus, tranzitív teljes gráf, úgy a $\Gamma_{R_c}(t_j)$ gráf (és ezzel együtt a t_i, t_j időintervallumbeli megismerési folyamat) akkor és csak akkor lesz az előzővel azonos tulajdonságú, ha

$$(2.11) \quad |E(t_j)| - |E(t_i)| = \frac{n_j(n_j - 1)}{2} - \frac{n_i(n_i - 1)}{2}.$$

Bizonyítás

Ha (2.11) teljesül, akkor az egyenlőséget átrendezve és $|E(t_i)| = n_i(n_i - 1)/2$ -t (a feltétel szerint) behelyettesítve adódik, hogy

$$(2.12) \quad |E(t_j)| = \frac{n_j(n_j - 1)}{2}.$$

Ez azt jelenti, hogy $\Gamma_{R_c}(t_j)$ teljes gráf. Továbbá az R_c reláció tulajdonságai miatt irányított, antiszimmetrikus, tranzitív gráf, így az 1. tétel alkalmazható $\Gamma_{R_c}(t_j)$ -re, ami azt jelenti, hogy a hozzá tartozó megismerési folyamat ideális és ezt akartuk belátni.

A tétel másik felének bizonyítása igen egyszerűen adódik, hiszen ha $\Gamma_{R_c}(t_j)$ ideális megismerési folyamat gráfja, akkor az 1. tétel alapján éleinek számára teljesül a (2.12) összefüggés, $\Gamma_{R_c}(t_i)$ -ről pedig feltettük, hogy éleinek számára teljesül az

$$(2.13) \quad |E(t_i)| = \frac{n_i(n_i - 1)}{2}$$

összefüggés. Ekkor az $|E(t_j)| - |E(t_i)|$ különbségre pontosan (2.11) adódik.

Ezzel a tételt beláttuk.

Mivel az ideális megismerési struktúrák között nem teszünk különbséget, azaz a megismerési folyamatban lévő vizsgálatok egyenrangúak és a 2. tétel szerint egy ideális megismerési folyamatot pontosan egy tranzitív, teljes gráf reprezentál, így egy aktuálisan megvalósuló megismerési struktúra és az ideális viszonya, a megfelelő gráfok élszámának arányával jól jellemezhető (pontosan ezt írja le a (2.9) összefüggés). Tehát a 3. tétel alapján $H(t_i, t_j)$ ténylegesen megfelel az 1., 2. követelményeknek.

A (2.9) egyenlőség jobb oldalának kifejtéséből adódik a $H(t_i, t_j)$ hatékonysági mutató egyszerűbb alakja:

$$(2.14) \quad H(t_i, t_j) = \frac{2(|E(t_j)| - |E(t_i)|)}{(2n_i + k(t_i, t_j) - 1) \cdot k(t_i, t_j)}.$$

A $H(t_i, t_j)$ mutató alsó és felső korlátairól a következőket mondhatjuk. A (2.10) összefüggést figyelembe véve kapjuk, hogy

$$(2.15) \quad 0 \leq E(t_i) \leq \frac{n_i(n_i - 1)}{2}$$

$$(2.16) \quad 0 \leq E(t_j) \leq \frac{n_j(n_j - 1)}{2}.$$

A (2.15) egyenlőtlenséget (2.16)-ból kivonva

$$(2.17) \quad 0 \leq |E(t_j)| - |E(t_i)| \leq \frac{n_j(n_j - 1)}{2} - \frac{n_i(n_i - 1)}{2}.$$

A (2.6), (2.7), (2.8) összefüggésekből következik, hogy

$$(2.18) \quad \forall i \geq 1 \Rightarrow k(t_i, t_j) \geq 1 \text{ és } n_i \geq 1.$$

Tehát (2.10) figyelembevételével adódik

$$(2.19) \quad \frac{n_j(n_j - 1)}{2} - \frac{n_i(n_i - 1)}{2} \geq 1.$$

Valamint

$$(2.20) \quad i = 0 \Rightarrow k(t_i, t_j) \geq 2 \text{ és } n_i = 0,$$

amely esetben szintén teljesül a (2.19) egyenlőtlenség, így a t_i, t_j időpontok bármely megválasztása mellett a (2.17) egyenlőtlenség leosztható a (2.19) egyenlőtlenség bal oldalával, amiből a $H(t_i, t_j)$ hatékonysági mutatóra az alábbi korlátok adódnak:

$$(2.21) \quad 0 \leq H(t_i, t_j) \leq 1.$$

A 3. tétel alapján világos, hogy $H(t_i, t_j) = 1$ hatékonyság pontosan akkor áll elő, ha az adott időintervallumban ideális a megismerési folyamat.

A $H(t_i, t_j) = 0$ hatékonyság pedig akkor, ha az adott időintervallumban végzett vizsgálatok egyáltalán nem kumulatívák (azaz teljesen elszigeteltek, függetlenek).

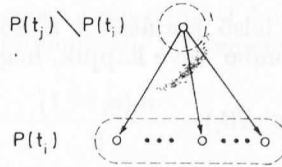
A teljes t_0, t_j időintervallum megismerési folyamatának hatékonyságára (2.7) alapján igen egyszerű összefüggés adódik:

$$(2.22) \quad H(t_0, t_j) = \frac{2 \cdot |E(t_j)|}{(k(t_0, t_j) - 1) \cdot k(t_0, t_j)} = \frac{2 \cdot |E(t_j)|}{(n_j - 1) n_j}.$$

Tegyük a kumulativitási struktúra és a megismerési folyamat hatékonyságának kapcsolatát szemléletesebbé!

Ehhez megmutatjuk a 3–6. ábrákon bemutatott elemi gráftípusokhoz tartozó hatékonyság értékeket. Hiszen, mint említettük, ezekből bármely megismerési folyamathoz tartozó kumulativitási Γ_{R_c} gráf felépíthető.

Először tekintsük az összegző típusú elemi struktúrát a t_i, t_j időintervallumban (lásd 11. ábra).

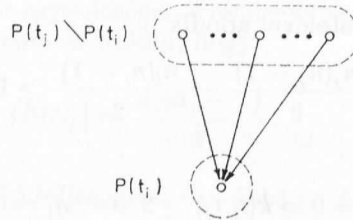


11. ábra

Ekkor a $k(t_i, t_j) = 1$ és $|E(t_j)| = |E(t_i)| + n_i$ összefüggések teljesülnek, amelyből adódik

$$(2.23) \quad H(t_i, t_j) = \frac{2(|E(t_i)| + n_i - |E(t_i)|)}{(2 \cdot n_i + 1 - 1) \cdot 1} = 1,$$

A kereső típusú elemi struktúra esetét a 12. ábra mutatja.



12. ábra

Ekkor az $n_i = 1$ és $|E(t_j)| = |E(t_i)| + k(t_i, t_j)$ összefüggésekből adódik

$$(2.24) \quad H(t_i, t_j) = \frac{2(|E(t_i)| + k(t_i, t_j) - |E(t_i)|)}{(2n_i + k(t_i, t_j) - 1) \cdot k(t_i, t_j)} = \\ = \frac{2}{2n_i + k(t_i, t_j) - 1} = \frac{2}{k(t_i, t_j) + 1},$$

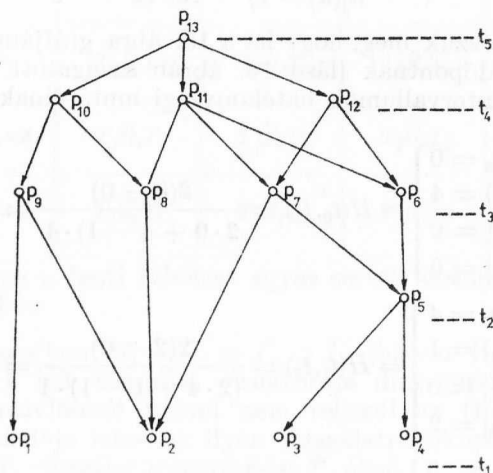
azaz, ha élünk a végtelen számú vizsgálat absztrakciójával, akkor

$$(2.25) \quad \lim_{k(t_i, t_j) \rightarrow \infty} H(t_i, t_j) = 0.$$

A (2.23) és (2.25) összefüggések igen jól mutatják, hogy a teljes megismerési folyamat hatékonyságát közelítőleg a kereső és összegző típusú részfolyamatok aránya határozza meg.

Ezek speciális eseteiként állnak elő az elemi ideális, illetve a széteső típusok (lásd 5., 6. ábra), amelyek hatékonyságára könnyen adódnak rendre a (2.23) és (2.25) összefüggések.

Szemléltetésként bemutatjuk egy fiktív megismerési struktúra $\Gamma_{R_0}(t_j)$ gráfjának $\Gamma'_{R_0}(t_j)$ ($j = 5$, lásd 13. ábra) parciális gráfját és megadjuk a $H(t_0, t_j)$ hatékonysági mutató értékét. Természetesen a $H(t_0, t_j)$ érték kiszámításánál a $\Gamma_{R_0}(t_j)$ összes éleinek számát vesszük figyelembe. Ezt a gráfot azonban nehezen áttekinthető volta miatt, csak szomszédossági mátrixával adjuk meg (lásd 14. ábra).



13. ábra

	1	2	3	4	5	6	7	8	9	10	11	12	13
1													
2													
3													
4													
5				1	1								
6				1	1	1							
7			1	1	1	1							
8			1										
9			1										
10		1	1						1	1			
11		1	1	1	1	1	1	1	1				
12		1	1	1	1	1	1	1					
13	1	1	1	1	1	1	1	1	1	1	1	1	1

14. ábra

d13.

Egy n szögpontú $G = (P, E)$ gráf szomszédossági mátrixa olyan $n \times n$ -es mátrix, amelynek bármely a_{ij} elemére fennáll a következő:

$$(2.26) \quad a_{ij} = \begin{cases} 1, & \text{ha } (p_i p_j) \in E \\ 0, & \text{ha } (p_i p_j) \notin E. \end{cases}$$

A szomszédossági mátrixból leszámítható, hogy

$$(2.27) \quad |P(t_j)| = n_j = 13 \quad \text{és} \quad |E(t_j)| = 41,$$

tehát példánk esetében

$$(2.28) \quad H(t_0, t_5) = \frac{2|E(t_j)|}{n_j(n_j - 1)} = \frac{82}{13 \cdot 12} \approx 0,52.$$

Érdekességként nézzük meg, hogy ha a 13. ábra gráfjának egyes szintjeit tekintjük egy-egy időpontnak (lásd 13. ábrán szaggatott vonallal jelezve), akkor az egyes időintervallumok hatékonysági mutatóinak sorozata hogyan alakul?

$$\left. \begin{array}{l} n_0 = 0 \\ k(t_0, t_1) = 4 \\ |E(t_0)| = 0 \\ |E(t_1)| = 0 \end{array} \right\} \Rightarrow H(t_0, t_1) = \frac{2(0 - 0)}{2 \cdot 0 + 4 - 1} \cdot 4 = 0$$

$$\left. \begin{array}{l} n_1 = 4 \\ k(t_1, t_2) = 1 \\ |E(t_1)| = 0 \\ |E(t_2)| = 2 \end{array} \right\} \Rightarrow H(t_1, t_2) = \frac{2(2 - 0)}{(2 \cdot 4 + 1 - 1) \cdot 1} = 0,5$$

$$\left. \begin{array}{l} n_2 = 5 \\ k(t_2, t_3) = 4 \\ |E(t_2)| = 2 \\ |E(t_3)| = 12 \end{array} \right\} \Rightarrow H(t_2, t_3) = \frac{2(12 - 2)}{(2 \cdot 5 + 4 - 1) \cdot 4} = 0,38$$

$$\left. \begin{array}{l} n_3 = 9 \\ k(t_3, t_4) = 3 \\ |E(t_3)| = 12 \\ |E(t_4)| = 29 \end{array} \right\} \Rightarrow H(t_3, t_4) = \frac{2(29 - 12)}{(2 \cdot 9 + 3 - 1) \cdot 3} = 0,56$$

$$\left. \begin{array}{l} n_4 = 12 \\ k(t_4, t_5) = 1 \\ |E(t_4)| = 29 \\ |E(t_5)| = 41 \end{array} \right\} \Rightarrow H(t_4, t_5) = \frac{2(41 - 29)}{(2 \cdot 12 + 1 - 1) \cdot 1} = 1.$$

3. Nem kumulatív vizsgálatpárok elemzése

A következőkben megvizsgáljuk, hogy mi a helyzet azokban az esetekben, amikor a vizsgálatpárok között nem áll fenn a kumulativitás (R_c) reláció.

Az előző pontokban bevezetett jelöléseket továbbra is megtartjuk, valamint jelöljük a T_i és T_j vizsgálatokban egyaránt felvett változók által F_i -ben, illetve F_j -ben kifesztett részgráfot rendre G_i , G_j -vel. (F_i , illetve F_j a T_i , illetve T_j vizsgálatok változó-struktúráját reprezentáló gráfok. Lásd az (1.9) összefüggést.)

Egy vizsgálat struktúráján a vizsgálatban felvett változók 1.2. definíció szerinti struktúráját értjük.

Ekkor az alábbi táblázat tartalmazza a T_i, T_j vizsgálatok struktúrájának összes lehetséges viszonyát.

A táblázatban \bar{R}_c jelöli az R_c reláció nem teljesülését.

A táblázatban szereplő számok csak a későbbi hivatkozást szolgálják.

	$V_i = V_j$	$V_i \subset V_j$	$V_i \cap V_j \neq \emptyset$	$V_i \cap V_j = \emptyset$
$G_i = G_j$	$T_j \bar{R}_c T_i$ ¹	$T_j R_c T_i$ ²	$T_j \bar{R}_c T_i$ ³	— ⁴
$G_i \cap G_j \neq \emptyset$	$T_j \bar{R}_c T_i$ ⁵	$T_j \bar{R}_c T_i$ ⁶	$T_j \bar{R}_c T_i$ ⁷	— ⁸
$G_i \cap G_j = \emptyset$	$T_j \bar{R}_c T_i$ ⁹	$T_j \bar{R}_c T_i$ ¹⁰	$T_j \bar{R}_c T_i$ ¹¹	$T_j \bar{R}_c T_i$ ¹²

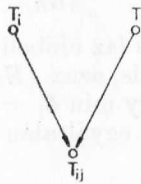
A következőkben a fenti táblázat egyes eseteit elemezzük, az azonosítási számok sorrendjében.

1. Ebben az esetben $G_i = G_j \Rightarrow \Gamma_i = \Gamma_j$ (hiszen $V_i = V_j \Rightarrow G_i = \Gamma_i$ és $G_j = \Gamma_j$). Tehát a két vizsgálat struktúrája megegyezik. Ekkor bár nem beszélünk kumulativitásról (mivel nem teljesül az (1.10) feltétel), mégis igen lényeges funkciója lehet az ilyen vizsgálatpároknak, hiszen ez az eset nem más, mint a T_i vizsgálat megerősítése T_j által.

Megjegyezzük azonban, hogy ezt az esetet tartalmazza a 2. eset, azaz a kumulativitás esete, így ez utóbbi sokkal gazdaságosabb (hatékonyabb), mivel a megerősítés és építés (további megismerés) egy lépésben történik.

2. Ez a kumulativitás esete, amellyel a 2. fejezetben részletesen foglalkoztunk.

3. Ebben az esetben nem teljesülnek a kumulativitás feltételei, azonban a két vizsgálat tartalmaz egy olyan közös részvizsgálatot, amelyre mindkettő kumulatív. Jelöljük e közös részvizsgálatot T_{ij} -vel.



15. ábra

Ugyanis legyen a T_{ij} részvizsgálat változóhalmaza $V_{ij} = V_i \cap V_j$, ekkor teljesülnek az alábbi összefüggések:

$$\left. \begin{array}{l} V_{ij} = V_i \cap V_j \Rightarrow V_{ij} \subset V_i \\ G_i = G_j \Rightarrow \Gamma_{ij} = G_i \subset \Gamma_i \end{array} \right\} \Rightarrow T_i R_c T_{ij}$$

$$\left. \begin{array}{l} V_{ij} = V_i \cap V_j \Rightarrow V_{ij} \subset V_j \\ G_i = G_j \Rightarrow \Gamma_{ij} = G_j \subset \Gamma_j \end{array} \right\} \Rightarrow T_j R_c T_{ij}$$

Ennek az esetnek a megismerési gráfját mutatja a 15. ábra.

Ez az eset tehát úgy fogható fel, mintha egy három vizsgálatból álló, kereső típusú vizsgálatsorozatot végeztünk volna el, ahol a T_{ij} vizsgálatot kumulálja a T_i és T_j vizsgálat is.

Ennek kapcsán kerül felszínre egy érdekes elemzési szempont, amelyet eddig nem említettünk. Ez pedig az, hogy az R_c kumulativitási reláció a különböző időpontokban végzett vizsgálatok halmazán definiál egy rendezést, amely nem feltétlenül esik egybe az időbeli rendezettséggel. Ez pontosan azt jelenti, hogy a kumulativitás definíciója „időtlen”, tehát nem zárja ki, hogy egy időben korábban végzett vizsgálat kumulatív legyen egy később végzett vizsgálatra nézve.

Így a megismerési struktúra egy új, számszerűen kifejezhető jellemzőjét kapjuk, ha azt vizsgáljuk, hogy a Γ_{R_c} gráf szögpontjainak megfelelő T_1, T_2, \dots, T_r vizsgálatok időbeli rendezettsége mennyire egyezik a kumulativitás relációja (R_c) által a vizsgálatok halmazán definiált rendezéssel.

Ennek vizsgálatához tegyük fel, hogy a T_1, T_2, \dots, T_r jelölésekben az indexek a szociológiai vizsgálatok időbeli sorrendjét jelölik.

3.1. Definíció

A T_1, T_2, \dots, T_r vizsgálatok halmazán értelmezett R_1 relációt *inverzió reláció*-nak nevezzük, ha

$$(3.1) \quad T_i R_1 T_j \Leftrightarrow T_i R_c T_j \text{ és } i < j \\ (1 \leq i \leq r, 1 \leq j \leq r).$$

Az R_c relációhoz hasonlóan az R_1 relációhoz is tartozik egy $\Gamma_{R_1} = (P_{R_1}, E_{R_1})$ irányított gráf. Ez utóbbi gráf éhalmaza pontosan azoknak a vizsgálatoknak megfelelő szögpontokat összekötő éleket tartalmazza, amely vizsgálatpároknál az időbeli sorrend nem egyezik meg a kumulativitási sorrenddel.

Képezzük tehát az alábbi δ_1 mutatót

$$(3.2) \quad \delta_1 = \frac{|E_{R_1}|}{|E_{R_c}|}.$$

Világos, hogy ha a két rendezés (az időbeli és a kumulativitási) megegyezik, akkor Γ_{R_1} éleinek száma minimális, azaz $|E_{R_1}| = 0$ és pontosan ekkor adódik δ_1 minimális értéke is, amelyre így $\min \delta_1 = 0$.

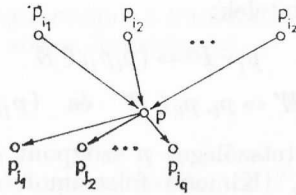
Ugyanakkor, ha a két rendezés egyáltalán nem egyezik meg, akkor $|E_{R_1}| = |E_{R_c}|$, azaz $\max \delta_1 = 1$, tehát

$$(3.3) \quad 0 \leq \delta_1 \leq 1.$$

Természetesen felmerül a kérdés, hogy milyen időpontot rendeljünk a megismerési (kumulativitási) struktúra azon vizsgálataihoz, amelyek (mint az a jelen 3. esetben is fennáll) két tényleges vizsgálat (a T_i és T_j vizsgálatok) kumulált közös részvizsgálataként adódnak. E probléma csak az inverzió reláció vizsgálatakor lép fel, ezért tárgyaljuk ezen a helyen. Itt egyetlen szempont érvényesítését tartjuk szükségesnek:

mivel saját időpont nélküli részvizsgálatokról van szó, ezért a tényleges vizsgálatok inverzióinak számát minimális mértékben növeljük csak a részvizsgálatokhoz rendelt időpontok.

Tekintsünk a megismerési struktúrát reprezentáló gráfban egy kumulált részvizsgálathoz tartozó p szögpontot és válasszuk ki p szomszédait (lásd 16. ábra).



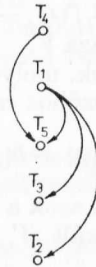
16. ábra

Ekkor a p szögpont által reprezentált vizsgálathoz a t_p időpontot rendeljük hozzá úgy, hogy a $\min(i_1, i_2, \dots, i_z) = \min(i)$ és $\max(j_1, j_2, \dots, j_q) = \max(j)$ rövidebb jelöléseket használva

$$(3.4) \quad t_p = \begin{cases} \min(i), & \text{ha } \min(i) \geq \max(j) \\ \max(j), & \text{ha } \min(i) < \max(j) \text{ és } z \leq q \\ \min(i), & \text{ha } \min(i) < \max(j) \text{ és } z > q. \end{cases}$$

Az így kapott t_p érték bármely p szögpont esetén (amely részvizsgálatot reprezentál) eleget tesz a kitűzött minimalitási kritériumnak.

Példaként az inverzió bemutatására tekintsük a 10. ábra megismerési struktúrájához tartozó Γ_{R_i} inverziós gráfot (ahol az időbeli rendezést a szögpontok indexei jelölik) és kiszámítjuk erre az esetre a δ_1 mutató értékét is (lásd 17. ábra).



17. ábra

Ebben az esetben tehát $|E_{R_i}| = 4$, $|E_{R_c}| = 10$, így $\delta_1 = 0.4$.

4. Ez az eset nem állhat elő, hiszen $V_i \cap V_j = 0$ miatt G_i és G_j részgráfok nem léteznek.

5. Ebben az esetben hasonló a helyzet, mint a 3. esetben, csak itt a közös részvizsgálat meghatározása nehezebb. Most megadjuk a T_{ij} közös részvizsgálat V_{ij} változóhalmazának kijelölésére szolgáló eljárást, illetve az ennek létezéséhez szükséges feltételeket. Mint ez a továbbiakból kiderül, ellentétben a 3. esettel, itt nem feltétlenül létezik a feltételeket kielégítő T_{ij} kumulált közös részvizsgálat.

3.2. Definíció

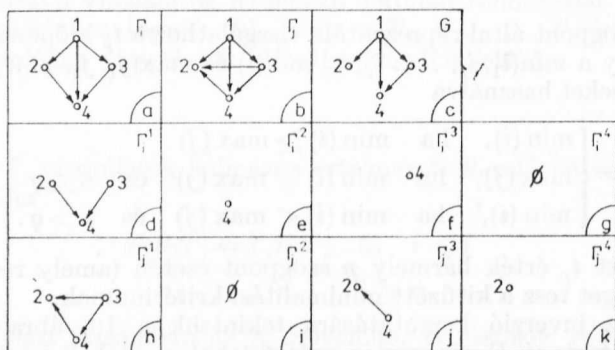
Legyen $\Gamma = (P, E)$ tetszőleges irányított gráf és $p_i \in P$.

A $\Gamma' = (P', E')$ gráfot a p_i szögpont által generált Γ -beli részgráfnak nevezzük, ha teljesülnek az alábbi feltételek:

$$(3.5) \quad p_j \in P' \Leftrightarrow (p_i p_j) \in E$$

$$(3.6) \quad (p_j p_k) \in E' \Leftrightarrow p_i, p_k \in P' \text{ és } (p_j p_k) \in E.$$

Jelöljük egy Γ gráfbeli tetszőleges p szögpont kimenő fokszámát $k(p)$ -vel, bemenő fokszámát $b(p)$ -vel. (Kimenő fokszámon a p -re illeszkedő kimenő, bemenő fokszámon a p -re illeszkedő bemenő élek számát értjük.)



18. ábra

Legyen tehát ebben az esetben $G_i \cap G_j = G$, valamint a G -t kifesztítő szögpontoknak megfelelő változók halmaza V .

V -be csak azok a változók kerülnek, melyekhez rendelt G -beli szögpontra teljesül az alábbi feltétel (p a v változóhoz rendelt G -beli szögpont):

$$(3.7) \quad v \in V \Leftrightarrow k(p) + b(p) \neq 0.$$

Ekkor a V_{ij} -beli változók pontosan azok a V -beli változók lesznek, amelyeket reprezentáló szögpontok által generált Γ_i , illetve Γ_j -beli részgráfok részei G -nek.

Formalizált leírással ez azt jelenti, hogy ha $p \in G$ és Γ_i^p , Γ_j^p rendre a p szögpont által generált Γ_i , illetve Γ_j -beli részgráfok, akkor a p szögpont által reprezentált v változókra az alábbi feltételnek kell teljesülnie:

$$(3.8) \quad v \in V_{ij} \Leftrightarrow \Gamma_i^p = \Gamma_j^p \subset G.$$

Ebből a feltételből következik, hogy az 5. esetben bár a $V \neq \emptyset$ feltételt biztosítjuk, mégis könnyen adódhat $V_{ij} = \emptyset$. Ennek illusztrálására a 18. $a-k$. ábrákon bemutatunk egy könnyen ellenőrizhető ilyen esetet. (Az ábrákon a szövegben használt jelölések láthatók.)

Felhívjuk a figyelmet a (3.8) összefüggés lényeges elméleti szerepére. Ez ugyanis azt fejezi ki, hogy két különböző vizsgálatban felvett változók nem

a nevük, hanem a vizsgálathoz tartozó megfelelő változó-struktúrában meglevő környezetük (generált részgráf) alapján tekinthetők azonosaknak. Ez egyben módot ad a bevezetésben leírt a ., c . problémák bizonyos mértékű kiküszöbölésére is.

A következő tételben megmutatjuk a kumulatív vizsgálatpár és a kumulált közös részvizsgálat közötti szoros összefüggést.

4. Tétel

Legyen T_i és T_j két szociológiai vizsgálat és $\Gamma_i = (P_i, E_i)$, $\Gamma_j = (P_j, E_j)$ a megfelelő változó-struktúrák gráfjai. Ha $T_j R_c T_i$ teljesül, akkor a (3.8) feltétel fennáll mindegyik V_i -beli változóra, azaz

$$(3.9) \quad T_j R_c T_i \Rightarrow V_{ij} = V_i.$$

(A jelölések megegyeznek a (3.8) összefüggésnél használtakkal.)

Bizonyítás

$$\begin{aligned} T_j R_c T_i \Rightarrow G_i = G_j = \Gamma_i \Rightarrow G = G_i \cap G_j = \Gamma_i \Rightarrow \forall v \in V_i : \Gamma_i^v = \\ = \Gamma_i^v \subseteq \Gamma_i = G \stackrel{(3.8)}{\Rightarrow} V_{ij} = V_i. \end{aligned}$$

Tehát a kumulált közös részvizsgálat egy „elfajult” esete az, amikor az egyik vizsgálat teljes egészében alkotja a közös részvizsgálatot.

6., 7. E két esetben azonos módon járunk el, mint az 5. esetben.

8. Itt szintén érvényesek a 4. esetre mondottak.

9., 10., 11., 12. Ezekben az esetekben a $G_i \cap G_j = \emptyset$ feltétel miatt, részvizsgálat erejéig sem beszélhetünk kumulativitásról. Így ezekben az esetekben a vizsgálatok teljes elszigetelődéséről beszélhetünk.

4. Kumulativitási struktúra elemzésének bemutatása ténylegesen elvégzett szociológiai vizsgálatokon

A következőkben négy elvégzett szociológiai vizsgálat felhasználásával „élesben” mutatjuk be az előző három fejezetben leírt modellezési és elemzési módszereket. A jelöléseket értelemszerűen az eddigieknek megfelelően használjuk, ezért ezek magyarázatára nem térünk külön ki.

Terjedelmi okokból itt nem a teljes változó-struktúrákat elemezzük. A T_1 vizsgálatból 8 változót, a T_2 vizsgálatból 9 változót, a T_3 vizsgálatból 7 változót és a T_4 vizsgálatból 4 változót tartottunk meg. Így az egyes vizsgálatokhoz (amelyeket a mobilitási jelenség vizsgálatára végeztünk) rendre az alábbi változó-struktúrák tartoznak.

A megfelelő $\Gamma_1, \Gamma_2, \Gamma_3, \Gamma_4$ gráfokat szomszédossági mátrixaikkal adjuk meg és a szögpontok számozása az eredeti számozásnak felel meg. A vizsgálatok elvégzésének időpontját (az inverzió elemzése miatt) feltüntettük a megfelelő szomszédossági mátrix mellett.

$$\Gamma_1 \text{ (1976)}$$

	1	2	4	5	7	8	10	35
1		1	1	1	1	1	1	
2				1	1	1	1	1
4	1	1		1	1	1	1	1
5								
7								
8								
10								
35			1	1	1	1		

$$\Gamma_2 \text{ (1977)}$$

	1	2	4	5	7	8	10	29	35
1									
2			1	1	1	1	1	1	1
4						1			
5		1	1	1		1	1	1	1
7									
8									
10									
29	1	1	1	1	1	1	1		1
35	1		1	1	1	1	1	1	

$$\Gamma_3 \text{ (1978)}$$

	2	5	7	10	32	35
2		1	1	1	1	1
5						
7						
10						
32						
35						

$$\Gamma_4 \text{ (1975)}$$

	5	7	32	35
5				1
7				1
32				
35				

Két szociológiai vizsgálat (T_i, T_j) maximális kumulált részvizsgálatának meghatározására a (3.8) összefüggés alapján az alábbi algoritmus adható:

1. Képezzük a $V_i \cap V_j = W$ változóhalmazt.
2. Képezzük a W halmaz elemei által Γ_i -ben kifeszített G_i gráfot.
3. Képezzük a W halmaz elemei által Γ_j -ben kifeszített G_j gráfot.
4. Képezzük a $G = G_i \cap G_j$ gráfot.
5. Hagyjuk el G -ből az izolált szögponthoz tartozó pontokat. Jelöljük az így megmaradó szögponthoz tartozó pontok által reprezentált változók halmazát V -vel.
6. Az összes megmaradt G -beli szögpontra szerkesszük meg külön-külön a Γ_i -beli generált részgráfját.
7. Az összes megmaradt G -beli szögpontra szerkesszük meg külön-külön a Γ_j -beli generált részgráfját.
8. Ha egy megmaradt G -beli p szögpontra megegyezik a 6. és 7. lépésben kapott két generált részgráf, akkor $p \in V_{ij}$.
9. Ha az összes megmaradt G -beli szögpontra a 8. összehasonlító lépést elvégeztük, akkor V_{ij} -ben megkaptuk a T_i és T_j vizsgálatok által kumulált maximális részvizsgálatot.

Megjegyezzük, hogy a 4. tétel alapján a fenti algoritmus elvégzésével, végeredményként a két vizsgálat kumulativitására is választ kapunk.

Terjedelmi okokból a fenti algoritmust csak a Γ_1 és Γ_2 gráfok összehasonlítására részletezzük, a teljes elemzésnek a végeredményét azonban közöljük.

A T_1, T_2 vizsgálatok maximális, kumulált részvizsgálatának meghatározása, a fenti algoritmus lépéseinek sorrendjében:

1. $W = \{1, 2, 4, 5, 7, 8, 10, 35\}$.

2.

		G_1							
		1	2	4	5	7	8	10	35
1			1	1	1	1	1	1	1
2					1	1	1	1	1
4			1	1		1	1	1	1
5									
7									
8									
10									
35					1	1	1	1	

3.

		G_2							
		1	2	4	5	7	8	10	35
1			1	1	1	1	1	1	1
2							1		
4									
5			1	1	1		1	1	1
7									
8									
10									
35		1		1	1	1	1	1	

4.

		G							
		1	2	4	5	7	8	10	35
1			1	1	1	1	1	1	
2							1		
4									
5									
7									
8									
10									
35					1	1	1	1	

5. Izolált szögpontok nincsenek, tehát

$$V = \{1, 2, 4, 5, 7, 8, 10, 35\}.$$

6.

		Γ_1^1					
		2	4	5	7	8	10
2				1	1	1	1
4		1		1	1	1	1
5							
7							
8							
10							

7.

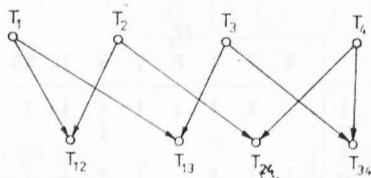
		Γ_2^1						
		2	4	5	7	8	10	35
2						1		
4								
5		1	1	1	1	1	1	1
7								
8								
10								
35			1	1	1	1	1	

8. Mivel $\Gamma_1^1 \neq \Gamma_1^2$, így az 1-es változó nem eleme a kumulált közös rész-vizsgálat V_{ij} változóhalmazának.

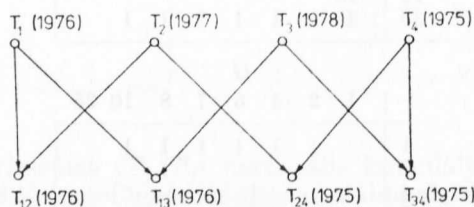
9. A többi V -beli változóra az előző 6., 7., 8. lépésekkel analóg módon járunk el, így ezek részletes közlését terjedelme miatt nem tartjuk célszerűnek.

Végeredményként: $V_{12} = \{7, 8, 10\}$

Az algoritmust az összes vizsgálatpárokra elvégezve, az alábbi megismerési (kumulatívítási) struktúra, azaz Γ_R gráf áll elő (lásd 19. ábra).



19. ábra



20. ábra

Végül megadjuk a teljes megismerési folyamat hatékonysági és inverziós mutatóját.

$$n_4 = 8, |E(t_4)| = 8, H(t_0, t_4) = \frac{2|E(t_4)|}{(n_j - 1)n_j} = 0.28.$$

A 20. ábrán az inverzió relációhoz (R_I) tartozó időpont hozzárendelésekkel (a (3.4) összefüggés alapján) mutatjuk be a Γ_R gráfot, amelyből kiderül, hogy az inverziós gráf Γ_{R_I} nem tartalmaz élt, azaz $|E_{R_I}| = 0$, így a $\delta_I = 0$ inverziós mutató érték adódik.

(Beérkezett: 1979. január 4-én.)

IRODALOMJEGYZÉK

1. BERGE, C.: Graphs and Hypergraphs North-Holland, 1973.
2. DÉNES, T.: Szekvenciális rendszerek gráfelméleti modellezése Információ Elektronika, 1978/6.
3. DÉNES, T.—GELLÉRI, P.: On the use of mathematics to sociology today In: Sociology of Science and Research Akadémiai Kiadó, megjelenés alatt.
4. KAUFMANN, A.: Pontok, élek, ívek... gráfok Műszaki Kiadó, Budapest, 1972.
5. KINDLER, J.—PAPP, O.: Komplex rendszerek vizsgálata Budapest, 1977. Műszaki Kiadó.
6. KÖNIG, D.: Theorie der endlichen und unendlichen graphen Leipzig, 1936. Akad. Verlagsges.

7. KUNSZT, Gy.: A tudományos kutatás logikai modellezése és tematikai irányítása Budapest, 1975. Akadémiai Kiadó.
8. O'MURCHEARTAIGH—PAYNE, C (eds.): Model fitting J. Wiley, 1977.
9. O'MURCHEARTAIGH—PAYNE, C. (eds.): Exploring data structures J. Wiley, 1977.
10. SREJDER, JU. A.: Egyenlőség, hasonlóság, rendezés Budapest, 1975. Gondolat Kiadó.
11. ZASZLAVSZKAJA, T. I.—RIVKINA, R. V.: A munkaerőmobilitás szociológiai kutatásának módszertani problémái Budapest, 1976. Közgazdasági és Jogi Könyvkiadó.
12. ZDRAVOMISZLOV, A.: A szociológiai kutatások módszertana Budapest, 1973. Kossuth Könyvkiadó.

GRAPH THEORETICAL TOOLS FOR THE STUDY OF THE CUMULATIVE BUILD-UP OF EMPIRICAL SOCIOLOGY

If sociological investigations are not considered as independent units but as certain phases of a cognitive process serving for description of a social phenomenon, the following problems will arise:

1. The set of the variables examined in two different investigations is usually not identical.

2. Representativeness of the various surveys is generally different, already on account of problem 1.

3. Also the measurement of variables considered identical is often different.

4. Widely different methods are applied for evaluation of the measurement data.

It is, therefore, a basic question, what is there to guarantee the commensurability of different analyses covering a given social phenomenon, or, in which way their cumulativeness can be described. In other words: do examinations bring us closer to the full cognition of the phenomenon?

In the paper the authors put forward a graph model that offers a solution of the problem. They give a mathematical definition of the cumulativeness of the cognitive process, of its structure, as well as of the structural description method of the contents of a variable. Thereby it becomes possible to select from an arbitrarily long sequence of sociological studies the variables the contents of which will be growing ever richer in the course of the cognition. The formalised character of the method enables also the calculation of the efficacy of the cognitive process of the given phenomenon, or, the trend of the efficacy. Finally, the authors set forth the algorithm of the method together with examples.

МЕТОДЫ ТЕОРИИ ГРАФОВ В ИССЛЕДОВАНИИ ИЕРАРХИЧЕСКОЙ СТРУКТУРЫ ЭМПИРИЧЕСКОЙ СОЦИОЛОГИИ

Если социологические исследования будут рассматриваться не в качестве самостоятельных подразделений, а как отдельные фазы познавательного процесса, призванного описывать какие-либо общественные явления, то возникают следующие основные проблемы:

1. Множество рассматриваемых переменных с точки зрения двух различных исследований чаще всего не совпадает.

2. Репрезентативность различных исследований — хотя бы вследствие проблемы № 1, чаще всего является различной.

3. Зачастую, не является одинаковым измерение переменных, которые считаются тождественными.

4. При оценке получаемых данных используются самые различные методы.

Таким образом основной вопрос кроется в том, что что обеспечивает сопоставимость различных исследований, касающихся рассматриваемых общественных явлений, т. е. как можно охарактеризовать их иерархию (кумулятивности)? Иными словами: приближают ли проводимые исследования к полноте познания какого-нибудь явления?

В своей работе авторы излагают такую модель в виде графа, которая предлагает средство для решения выдвигаемых вопросов. Авторы математически дают определение кумулятивности процесса познания, его структуры, а также порядок структурного описания содержания какой-либо переменной. Благодаря этому становится возможным, чтобы из серии любых по продолжительности социологических исследований можно было выбирать те переменные, содержание которых в ходе процесса познания все более обогащается. Формальность их метода позволяет рассчитывать эффективность познания данного явления, т. е. тренд этой эффективности. В заключении приводится иллюстрированный примерами алгоритм данного метода.