OM BRUG AF SANDSYNLIGHEDS-REGNING TIL VURDERING AF VARIGHEDEN AF ET PROJEKT

MED TILFØJELSE JANUAR 2004

Niels Herman Hansen

TECHNICAL REPORT
IMM-REP-1998-21

IMM

I kapitel 5 af bogen 'Planlægning – i en foranderlig verden' beskriver Steen Lichtenberg en metode til vurdering af varigheden af et projekt, som jeg finder problematisk. Jeg har derfor taget problemerne vedrørende vurdering af varigheden af et projekt op til fornyet overvejelse, og formålet med det følgende er at give en kort fremstilling af mine overvejelser. En mere detaljeret fremstilling kan findes i [1].

1. Den til grund liggende logiske model af et projekt

Som det fremgår af det just nævnte kapitel 5 i Steen Lichtenbergs bog er det anerkendt praksis at beskrive et projekt ved brug af et netværk. Det falder uden for de her givne rammer at diskutere det. Vi nøjes med at se på følgende eksempel:

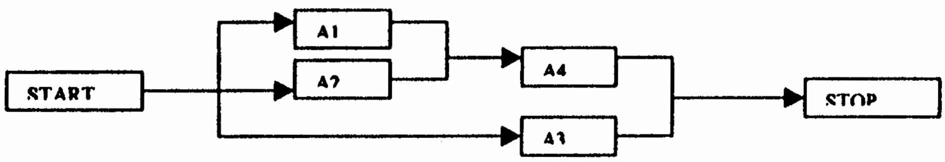

Fig1: Netværk der beskriver et hypotetisk projekt som vi benævner 'HPR'.

Fig 1 viser at HPR omfatter 4 aktiviteter A_1, A_2, A_3 og A_4 som skal udføres i overensstemmelse med følgende princip:

I: A_1, A_2 og A_3 påbegyndes når HPR startes

II: A_4 påbegyndes når A_1 og A_2 begge er afsluttet

III: HPR er slut når A_3 og A_4 begge er afsluttet

Når alt kommer til alt er et projekt en aktivitet sammensat af aktiviteter, der er sammensat af aktiviteter. Med det i tankerne skal vi nu se, hvorledes man ved en trinvis procedure kan opdele HPR i aktivitetspar, der skal udføres 'parallelt' eller 'i serie'. Symbolet '≡' som anvendes i fig. 2 betegner (som sædvanlig) ækvivalens, og det er underforstået at ækvivalente aktiviteter/aktivitetspar kan erstatte hinanden. Dvs. hvis vi tager de tre trin i fig. 2 i omvendt rækkefølge kommer vi tilbage til fig 1.

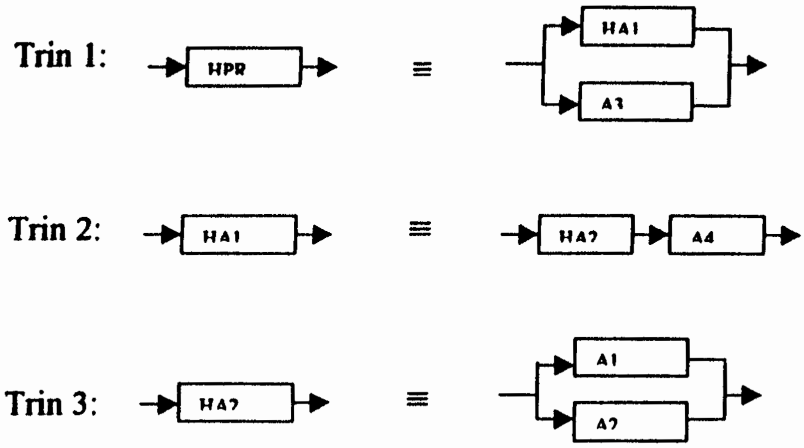

Fig 2: Opsplitning af HPR i aktivitetspar, der skal udføres 'parallelt' eller 'i serie'

2. Om varigheden af et par af aktiviteter, der skal udføres 'parallelt' eller 'i serie'

Som udgangspunkt tager vi følgende definition:

D: Givet A betegner en aktivitet betegner BT(A): begyndelsestidspunktet for A; ST(A): sluttidspunktet for A og VH(A): varigheden af A

Af denne definition følger det trivielt at:

$$T1: VH(A) = ST(A) - BT(A)$$

Vi indfører nu aktiviteterne PA1, PA2 og PA således:

Fig. 3: Ækvivalens mellem to aktiviteter PA1 og PA2, der skal udføres 'parallelt' og en enkelt aktivitet PA.

To aktiviteter, der skal udføres 'parallelt' udføres samtidigt, og det er underforstået, at to aktiviteter der udføres samtidigt også startes samtidigt, dvs.:

$$DPA1: BT(PA) = BT(PA1) = BT(PA2)$$

Det er endvidere underforstået, at en aktivitet bestående af to aktiviteter der udføres samtidigt ikke er afsluttet før de begge er afsluttet, dvs.:

$$DPA2: ST(PA) = \max \{ST(PA1), ST(PA2)\}$$

DPA1 og DPA2 udgør i forening en definition af den ækvivalens der er postuleret i fig 3, og det følger af T1 og DPA2 at:

$$VH(PA) = ST(PA) - BT(PA)$$
$$= \max\{VH(PA1) + BT(PA1), VH(PA2) + BT(PA2)\} - BT(PA). \qquad (1)$$

Da det for vilkårlige reelle tal x, y og c gælder at

$$RR1: \max\{x,y\} - c = \max\{x-c, y-c\}$$

følger det af (1) og DPA1 at:

$$TPA: VH(PA) = \max\{VH(PA1), VH(PA2)\}.$$

Vi indfører dernæst aktiviteterne SA1, SA2 og SA således:

Fig. 4.: Ækvivalens mellem to aktiviteter SA1 og SA2, der skal udføres 'i serie' og en enkelt aktivitet SA.

En aktivitet bestående af to aktiviteter der skal udføres 'i serie' påbegyndes når den første påbegyndes og sluttes når den sidste afsluttes, dvs.:

$$DSA1: BT(SA) = BT(SA1) \wedge ST(SA) = ST(SA2)$$

Når to aktiviteter skal udføres 'i serie' kan den sidste ikke påbegyndes før den første er afsluttet, men det er underforstået at den påbegyndes når den første afsluttes dvs.:

$$DSA2: BT(SA2) = ST(SA1)$$

DSA1 og DSA2 udgør i forening en definition af den ækvivalens der er postuleret i fig 4, og det følger ved brug af T1, DSA2 og DSA1 at

$$VH(SA) = ST(SA) - BT(SA) = ST(SA2) - BT(SA1)$$
$$= (STA(SA2) - BT(SA2)) + (ST(SA1) - BT(SA1))$$

hvilket i forening med T1 giver:

$$TSA: VH(SA) = VH(SA1) + VH(SA2)$$

3: Algoritme til beregning af varigheden af et projekt

Med henvisning til fig. 2 kan vi nu konstatere

$$\text{Trin 1} \wedge TPA \Rightarrow VH(HPR) = \max\{VH(HA1), VH(A3)\}$$
$$\text{Trin 2} \wedge TSA \Rightarrow VH(HA1) = VH(HA2) + VH(A4)$$
$$\text{Trin 3} \wedge TPA \Rightarrow VH(HA2) = \max\{VH(A1), VH(A2)\}$$

For at forenkle skrivearbejdet definerer vi

$$D1:\ V_1 = VH(A1)\ ;\ V_2 = VH(A2)\ ;\ V_3 = VH(A3)\ ;\ V_4 = VH(A4)$$

Med dette og det indledende resultat som udgangspunkt konstruerer vi nu følgende algoritme ved at tage de tre trin i omvendt rækkefølge

$$AVH:\ h_1 := \max\{V_1, V_2\};\ h_2 := h_1 + V_4\ ;\ VH(HPR) := \max\{h_2, V_3\}$$

Symbolet ':=' som optræder i AVH er lånt fra programmeringssprogene ALGOL og PASCAL hvor 'a: = b' betyder 'sæt a lig med b'. Når ordren 'a:=b' er udført har vi a=b, dvs. når vi har udført de beregninger der er foreskrevet ved hjælp af algoritmen AVH har vi

$$RVH:\quad VH(HPR) = \max\{h_2, V_3\} = \max\{h_1 + V_4, V_3\}$$
$$= \max\{\max\{V_1, V_2\} + V_4, V_3\}$$

Dette er grundlaget for definitionen D2 i afsnit 5.

4. Om varigheden af en aktivitet

Begyndelsestidspunktet for en aktivitet falder ikke nødvendigvis sammen med begyndelsestidspunktet for en arbejdsdag og sluttidspunktet for en aktivitet falder ikke nødvendigvis sammen med sluttidspunktet for en arbejdsdag. Hertil kommer at alle arbejdsdage ikke har samme længde. Trods dette er det ikke desto mindre naturligt at forudsætte at varigheden af en aktivitet er et helt antal arbejdsdage. Kort sagt: Det forudsættes, at det for en vilkårlig aktivitet A gælder at:

$$GF1:\quad VH(A) \in N^*$$

hvor 'N*' betegner mængden af hele positive tal. Hertil kan så føjes: Jeg har valgt tidsenheden 1 dag for det er den naturlige mindste tidsenhed, når vi taler om aktiviteter, der indgår i et projekt. Det er imidlertid ikke udelukket at vælge en anden tidsenhed – f.eks. 1 uge eller 1 måned.

GF1 fastsætter ikke nogen øvre grænser for varigheden af en aktivitet, men i praksis vil der være en øvre grænse. Det er f.eks. meningsløst at forestille sig at en aktivitet kan vare 1 million dage. Med det som udgangspunkt vil vi videre forudsætte, at det for enhver aktivitet A gælder at:

$$GF2:\quad \exists\, svh \in N^*\ ;\ VH(A) \leq svh$$

Vi forudsætter endvidere, at det for en vilkårlig aktivitet A gælder at

$$GF3:\ \exists\, mvh \in N^*\ ;\ VH(A) \geq mvh$$

Idet vi bruger symbolet '$\{m : M\}$' som betegnelse for mængden $\{n \in N^* \mid m \leq n \leq M\}$ kan forudsætningerne GF1, GF2 og GF3 sammenfattes i følgende forudsætning

$$GF:\ \exists\, (mvh\ ,\ svh) \in N^* \times N^*\ ;\ VH(A) \in \{mvh : svh\}$$

I overensstemmelse med dette forudsætter vi nu

$$F: \forall i \in \{1:4\};\ \exists (g_i, G_i) \in N^* \times N^*;\ V_i \in \{g_i : G_i\}$$

5. Om største og mindste varighed af et projekt

Som udgangspunkt har vi:

$$\text{RR2:}\quad \min(\max\{x,y\}) = \max\{\min(x), \min(y)\} \wedge$$
$$\max(\max\{x,y\}) = \max\{\max(x), \max(y)\}$$
$$\text{RR3:}\quad \min(x+y) = \min(x) + \min(y) \wedge \max(x+y) = \max(x) + \max(y)$$

Hvis vi nu læser algoritmen AVH bagfra og benytter RR2 og RR3 samt F får vi

$$\max(VH(HPR)) = \max\{\max(h_2), G_3\} \wedge \min(VH(HPR)) = \max\{\min(h_2), g_3\}$$
$$\max(h_2) = \max(h_1) + G_4 \wedge \min(h_2) = \min(h_1) + g_4$$
$$\max(h_1) = \max\{G_1, G_2\} \wedge \min(h_1) = \max\{g_1, g_2\}$$

Hvis vi læser dette fra neden opad ledes vi til følgende algoritmer

$$\text{ASVH}: h_1 := \max\{G_1, G_2\};\ h_2 := h_1 + G_4;\ G := \max\{h_2, G_3\}$$
$$\text{AMVH}: h_1 := \max\{g_1, g_2\};\ h_2 := h_1 + g_4;\ g := \max\{h_2, g_3\}$$

Heraf ser vi at hvis vi beregner G ved hjælp af algoritmen ASVH og g ved hjælp af algoritmen AMVH så får vi VH(HPR) $\le$ G og VH(HPR) $\ge$ g. Givet dette og RVH kan funktionen VHPR defineres således:

$$\text{D2:}\quad \text{VHPR}: \{g_1 : G_1\} \times \{g_2 : G_2\} \times \{g_3 : G_3\} \times \{g_4 : G_4\} \rightarrow \{g : G\}$$
$$\text{VHPR}(V_1, V_2, V_3, V_4) = \max\{\max\{V_1, V_2\} + V_4, V_3\}$$

Hermed er det udtrykt formelt at varigheden af HPR er en funktion af varigheden af aktiviteterne A1, ... , A4.

6. Om sandsynligheden for at VH(HPR) = k

Set fra en sandsynlighedsteoretisk synsvinkel er VH(HPR) en stokastisk variabel, dvs. en funktion defineret på et sandsynlighedsfelt (Ω, Á, P) (jfr. afsnit 0.2.1 i [3]). I den her givne sammenhæng vil vi med D2 i tankerne definere udfaldsrummet Ω således:

$$\text{D3:}\quad \Omega = \{g_1 : G_1\} \times \{g_2 : G_2\} \times \{g_3 : G_3\} \times \{g_4 : G_4\}$$

Á, som generelt skal være en σ - algebra af delmængder af Ω, kan i den her givne sammenhæng identificeres med mængden af alle delmængder af Ω. Elementerne i Á repræsenterer hændelser, og vi vil se på de hændelser der repræsenteres af følgende elementer i Á.

$$D4: \quad \forall k \in \{g : G\} \; ; H_k = \{\omega \in \Omega \mid VHPR(\omega) = k\}$$

P som kaldes 'et sandsynlighedsmål', er pr definition en afbildning Á $\rightarrow$ [0, 1] som skal opfylde visse betingelser, som i den givne sammenhæng medfører at

$$P(\Omega) = \sum_{\omega \in \Omega} P(\{\omega\}) = 1 \tag{1}$$

$$P(H_k) = \sum_{\omega \in Hk} P(\{\omega\}) \tag{2}$$

For fuldstændighedens skyld tilføjes det, at vi i overensstemmelse med D4 har

D5 : '$P(H_k)$' betegner sandsynligheden for at hændelsen VH(HPR) = k indtræffer hvis HPR udføres.

For at komme videre må vi tillægge størrelserne $P(\{\omega\})$ numeriske værdier, og det leder til følgende postulat:

$$EQP: \quad \exists \; p \in \,]0, 1[\; ; \; \forall \omega \in \Omega \; ; P(\{\omega\}) = p$$

Idet '#M' betegner antallet af elementer i mængden M følger det af (1) at

$$EQP \Rightarrow \forall \omega \in \Omega \; ; P(\{\omega\}) = (\#\Omega)^{-1}$$

hvilket i forening med (2) giver

$$PVH: \; EQP \Rightarrow \forall k \in \{g : G\} \; ; P(H_k) = \#H_k / \#\Omega$$

7. To forskellige begrundelser for påstanden EQP

Betydningen af resultatet PVH ligger i at det knytter en forbindelse mellem den moderne matematiske definition af sandsynlighed og den såkaldt 'klassiske' definition af sandsynlighed, som Laplace formulerer således på s. 6 i [4]:

> »The theory of chance consists in reducing all the events of the same kind to a certain number of cases equally possible, that is to say, to such as we may be equally undecided about in regard to their existence, and in determining the number of cases favorable to the event whose probability is sought. The ratio of this number to that of all cases possible is the measure of this probability, which is thus simply a fraction whose numerator is the number of favorable cases and whose denominator is the number of all the cases possible.«

Laplace bruger ikke begrebet 'udfaldsrum', for det var endnu ikke 'opfundet' da han skrev sit essay, men det giver mening at reformulere hans synspunkt således:

$$EQP1 \Rightarrow \forall \omega \in \Omega \; ; P(\{\omega\}) = (\#\Omega)^{-1}$$

EQP1 : Elementerne i Ω repræsenterer »a certain number of cases equally possible, that is to say [cases] we may be equally undecided about in regard to their existence«.

Det er ikke uden videre givet hvordan Laplace's forklaring af udtrykket »equally possible« skal tolkes, men der kan ikke være tvivl om at

$$BG1: \quad EQP1 \Rightarrow EQP$$

Sagt med andre ord: Laplace's definition af begrebet sandsynlighed giver os en begrundelse for påstanden EQP. Denne definition er siden hen blevet kritiseret, og von Mises har foreslået en alternativ definition som leder til følgende begrundelse for EQP formuleret ved delvis genbrug af Laplace's ord

$$BG2: \quad EQP2 \Rightarrow EQP$$

EQP2: elementerne i Ω repræsenterer »a certain number of cases equally [probable], that is to say [cases which will occur equally frequent if HPR is performed a great number of times]«.

8. Diskussion af et numerisk eksempel

Lad os nu se på følgende numeriske eksempel

$$EHPR: \quad \Omega = \{6:10\} \times \{6:10\} \times \{14:18\} \times \{6:10\}$$

Dette medfører først at: $\# \Omega = 5^4 = 625$, og ved brug af algoritmerne ASVH og AMVH får vi G =20 og g = 14. De numeriske værdier af størrelserne $\#H_k$, som indgår i SVH findes ved brug af følgende algoritme skrevet i programsproget MATLAB

```
AM:  am = zeros(1,7) ;
     for v1 = 6:10   for v2 = 6:10   for v3 = 14:18   for v4 = 6:10
     h1 = max (v1, v2) ; h2 = h1 + v4 ; h3 = max(h2, v3) -13 ;
     am(h3) = am(h3) +1 ; end, end, end, end
```

k:	14	15	16	17	18	19	20
$\#H_k$:	14	46	105	151	184	80	45

Tabel 1: Resultat af brugen af AM. $\#H_k$ = am (k − 13)

Hvis vi nu godtager påstanden EQP følger det af PVH at vi f.eks. har:

$$P(H_{18}) = 184/625 \tag{1}$$

men hvad betyder det? Ifølge den såkaldte 'frekvensfortolkning' af det matematiske begreb sandsynlighed betyder det at hændelsen VH(HPR) = 18 vil indtræffe rundt regnet 184 gange hvis HPR udføres 625 gange. Dette er 'klar tale', men er det egentlig 'realistisk' at forestille sig at HPR kan udføres 625 gange? Et projekt er - når alt kommer til alt – 'en engangsforeteelse'. Dette antyder at (1) ikke har en 'realistisk' frekvensfortolkning, og det sætter spørgsmålstegn ved den begrundelse for påstanden EQP som vi benævnte BG2.

Dette leder til den alternative begrundelse BG1. Det er som allerede nævnt ikke umiddelbart klart hvordan Laplace's forklaring af udtrykket »equally possible« skal forstås, men jeg foreslår følgende fortolkning af BG1:

> FTBG1: Påstanden EQP er sand SÅFREMT der ikke er grund til at tro, at eet snarere end et andet af de »cases« der repræsenteres af elementerne i Ω vil blive realiseret hvis HPR bliver udført.

Pointen i denne fortolkning er, at udtrykket 'grund til at tro' bringer den såkaldte 'degree of belief - fortolkning' af det matematiske begreb sandsynlighed ind i billedet. Det falder uden for de her givne rammer at forklare dette. Jeg nøjes med at bemærke, at de indledende beregninger i forening med denne fortolkning kan begrunde følgende påstand:

> SVH : det er muligt det vil tage mindre end 16 tidsenheder eller mere end 18 tidsenheder at udføre HPR, men det er 'mere sandsynligt' i betydningen 'mere troværdigt' at det vil tage 16, 17 eller 18 tidsenheder at udføre HPR.

Dette forudsætter at vi godtager påstanden EQP. Det bør derfor tilføjes, at jeg i [1] også opererer med forudsætninger, der er uforenelige med påstanden EQP. Det vil jeg dog ikke komme ind på fordi det fundamentale problem er, at resultatet af en sandsynlighedsberegning er uden praktisk værdi med mindre det har en realistisk fortolkning. Det kan diskuteres om SVH er en realistisk fortolkning af de indledende resultater, men det falder uden for de her givne rammer at gøre det.

9. Afsluttende bemærkninger

Jeg startede med den påstand, at Lichtenbergs metode til vurdering af varigheden af et projekt er problematisk. Det er i sig selv problematisk, for der er mange der mener, at denne metode har vist sin anvendelighed i praksis. Pladsen tillader ikke en diskussion af dette, men det kan være på sin plads at komme med nogle afsluttende bemærkninger.

Jeg vil ikke påstå at min metode – dvs. den her beskrevne – er bedre i den forstand, at den er mere 'virkelighedstro' end Lichtenbergs, men jeg vil samtidig påstå at Lichtenbergs metode heller ikke er mere 'virkelighedstro' end min. Jeg vil imidlertid påstå at min metode er bedre, fordi den i modsætning til Lichtenbergs ikke er baseret på brug af intuitivt begrundede approximationer. De resultater jeg når frem til følger med logisk nødvendighed af de anførte definitioner. Herved opnår jeg samtidig, at de fundamentale fortolkningsproblemer, som hyppigt diskuteres, kan præciseres. Hertil kommer at min metode kan udmøntes i enkle beregningsprocedurer, som kan realiseres ved brug af programmeringssproget MATLAB. Det er kun demonstreret ved et enkelt eksempel i afsnit 8, men interesserede kan finde mere i [1].

Tilbage står den kendsgerning at dette ikke er afprøvet i praksis, men jeg har aftalt med cand. polit. Henrik Kruckow at vi i fællesskab skal afprøve min metode.

Referencer:

[1] Niels Herman Hansen: *Notat vedr. brug af begrebet sandsynlighed i tidsplanlægning.* IMM Research report, som forventes afsluttet jan. 1999.

[2] Steen Lichtenberg: *Projektplanlægning - i en foranderlig verden.* Polyteknisk forlag, 3. udg. 1990.

[3] Knut Conradsen: *En introduktion til statistik.* IMM- DTU, 6. udg. 1995.

[4] Pierre Simon, Marquis de Laplace: *A Philosophical Essay on Probabilities.* Dover Publications S166, 1951.

TLFØJELSE JANUAR 2004

Det nævnte notat [1] foreligger dags dato kun i kladde, og den foreslåede metode er ikke blevet afprøvet i praksis. Det skyldes to ting:

1) KUN NOGLE projekter kan på den her angivne måde opsplittes i aktivitetspar, der skal udføres 'parallelt' eller 'i serie'.
2) Der er i mellemtiden kommet nye artikler om Stochastic PERT Networks, som bør tages i betragtning.

Givet dette har denne rapport følgende to PRINCIPIELLE formål:

1) At vise hvorledes den logiske struktur af et problem kan danne basis for en sandsynlighedsberegning.
2) At påpege, at det kan diskuteres hvorledes den beregnede sandsynlighed skal fortolkes.