

# Discussão de alguns métodos de análise de seqüências comportamentais

PAULO GAMA (\*)

*«Em poucas áreas de investigação biológica é tão fácil obter resultados e tão difícil de os explicar como no caso do estudo do comportamento animal.»*

(Simpson, 1969)

A análise de seqüências tornou-se já um método popular, frequentemente empregue na descrição e interpretação do comportamento animal. Utiliza-se para analisar dependências sequenciais ou transições entre dois ou mais actos em situações de agressão, corte, jogo, actividades de limpeza e também em estudos de comunicação. A par de outros procedimentos quantitativos, o seu uso torna possível obter maior grau de concisão nos resultados e na interpretação das observações bem como o tratamento de uma maior quantidade de dados que com métodos essencialmente qualitativos e descritivos se tornava impraticável. A estatística e mais recentemente a elaboração de modelos descritivos e a utilização de simulação em computadores, tornaram-se instrumentos fundamentais de análise em qualquer estudo etológico. É evidente que a quantificação, só por si, não corresponde a uma maior objectividade ou cientificidade da Etologia, ou de qualquer outra disciplina científica. No entanto, quantificar permite o tratamento de maior número de dados, a construção de imagens mais claras do que se observa, e a assentar em critérios que são geralmente

aceites para confirmar ou eliminar hipóteses, o que permite reduzir a relatividade das interpretações. Mas nem a Etologia é mais científica por usar um maior número de métodos quantitativos, nem deve ser confundida quantificação com objectividade. A subjectividade do observador continua presente na recolha da informação e não é a utilização de procedimentos estatísticos, mais ou menos complexos, para analisar essa informação que a vai remover.

A análise estatística é, porém, um instrumento extremamente importante no processo de confirmação ou eliminação de hipóteses e na resposta às questões de investigação formuladas. Permite ordenar e estruturar os dados de modos que não são aparentes ou nítidos numa observação dos dados sem manipulação, mas que constituem informação contida neles. Os sistemas biológicos e particularmente o comportamento podem ser muito complexos e a análise estatística é essencial para que se possa entender o que está a decorrer. Embora, por vezes, se utilizem procedimentos estatísticos excessivamente complexos como substituto de clareza de ideias ou de um bom delineamento experimental. Por isso, deve procurar-se adaptar o método de análise ao animal e não o animal ao método (Bekoff, 1977).

---

(\*) Assistente na Universidade de Coimbra.

Em inúmeras situações é particularmente a sequência pela qual os comportamentos ocorrem e não as frequências ou a duração dos mesmos que importa analisar. Por exemplo, o estudo da corte numa ave implica o registo da ordem pela qual os comportamentos se sucedem para se avaliar do grau de padronização da corte, ou das fases mais decisivas do processo. Também, comportamentos que tendem a ocorrer associados ou se apresentam pela mesma ordem poderão indicar que existe um mecanismo causal que lhes é comum. O estudo dos modelos de causalidade beneficiou bastante com a utilização da análise de sequências que possibilitou o exame de associações entre actos de uma forma mais objectiva (Slater, 1973). A análise de sequências revela-se bastante útil em estudos sobre interacções entre indivíduos, quando se pretende analisar a forma como essas interacções evoluem. O conhecimento da estrutura de um combate, dos pontos de maior incerteza ou de maior definição, podem ser aproximados através de análise de sequências. Há, no entanto, um certo número de problemas e limitações que serão expostos, bem como algumas formas de os contornar.

#### O QUE SÃO UNIDADES DE COMPORTAMENTO?

Os animais apresentam formas de comportamento que têm um certo padrão sequencial ou de repetição que pode ser mais ou menos evidente, mais ou menos variável e que, embora nalguns casos possa ter uma tipificação bastante elevada, está usualmente dependente de estímulos, internos e externos, susceptíveis de actuar durante o desenrolar do comportamento no sentido de modificar o seu curso. A descrição da locomoção de um gato não é feita descrevendo a posição relativa das diferentes patas ao longo da progressão, os ângulos que formam, etc. Há uma consistência bastante grande na forma de actuação dos 4 membros e do gato em si

quando se desloca. Essa consistência não envolve uma constância absoluta. Ocorrem variações, que, no entanto não colidem com a imagem de regularidade do conjunto.

Uma sequência numérica do género 1, 3, 6, 10, 15, 21, 28, 36, ou 1, 2, 3, 2, 4, 5, 6, 5, 7, 8, 9, 8, tem um padrão interno que permite, depois de conhecido, saber a sucessão seguinte de números. Trata-se de um padrão fixo, totalmente predizível. No caso do gato só é possível prever que um determinado movimento poderá suceder a outro com uma certa probabilidade. Tratando-se de uma probabilidade muito elevada, como no exemplo da locomoção, é possível considerar todo o processo como uma unidade. O nosso cérebro é particularmente eficiente a captar regularidades no mundo real — a constância dos fenómenos é essencial para a programação de acções — pelo que a mera observação cuidada dos comportamentos de um animal torna possível individualizar «unidades» de comportamento, com base nessas regularidades, de uma forma completamente intuitiva (Dawkins, 1983).

Se elaborarmos um programa para correr num computador constituído por uma série de regras para gerar uma sequência de números, de letras ou de proposições, e nele incluímos uma expressão que gere um valor aleatoriamente, com uma certa frequência, que também pode ser variável, sendo esse valor incluído na sequência, estamos em condições de gerar sequências que não são inteiramente lógicas porque incluem um elemento de aleatoriedade. A sequência será tanto menos 'lógica' quanto mais frequentes forem esses elementos imprevisíveis. Nestas circunstâncias já não será possível prever os elementos que se seguem numa sequência incompleta (como os exemplos acima) com certeza; apenas com uma determinada probabilidade (Dawkins, 1983). É possível afirmar que A sucede a B com uma probabilidade elevada P, mas não se pode afirmar que sempre que ocorre B temos A a seguir. Se agora nos colocarmos na posição inversa

confrontados com uma sequência gerada por um programa destes, mas do qual não conhecemos as regras, o nosso problema será o de encontrar as regularidades e constâncias que resultam das regras do programa e, eventualmente, descobri-las. É precisamente para isso que a análise de sequências é extremamente útil, na pesquisa de padrões de associação entre os elementos. Mas a analogia entre estes programas e o comportamento animal acaba aqui, porque o «programa» destes é muito mais vasto, não tem regras simples, salvo raras exceções, é aberto e está dependente de inúmeros factores, internos e externos, que condicionam ou dirigem a ordem da «sequência» comportamental. O exemplo, no entanto, serve, como toda a redução, para compreendermos o princípio em que se baseia este tipo de análise.

Antes de passarmos aos métodos há ainda aspectos relacionados com a definição das «unidades de comportamento» e com os processos de registo, com implicações na recolha dos dados, que é necessário referir.

#### DEFINIÇÃO DAS UNIDADES DE COMPORTAMENTO

O estabelecimento de unidades discretas, bem definidas, que o investigador extrai de um *continuum* de movimentos e contracções musculares que é o comportamento de um animal, constitui o primeiro problema que se coloca a uma análise deste tipo. Em alguns casos essas unidades parecem naturalmente individualizadas, enquanto noutros se torne bastante difícil proceder a uma divisão artificial do comportamento. A própria escolha das categorias depende muito mais do tipo de questões colocadas que de qualquer característica inerente ao comportamento. Em última análise, a sua formulação reflecte as preocupações e objectivos do investigador (Martin & Bateson, 1986).

As «unidades de comportamento» têm que ser discretas e definidas com precisão, de modo que seja possível distinguir o fim de

uma e o começo de outra (pode ser difícil distinguir dois actos de investigação contíguos). Devem ser independentes umas das outras para que não representem apenas formas distintas de medir a mesma coisa. O catálogo deve ser vasto o suficiente para cobrir toda a actividade do animal sem hiatos e permitir responder às questões formuladas, mas categorias a mais também significam transições muito pouco frequentes o que origina sérios problemas de análise. Também uma particularização excessiva de certos comportamentos e não de outros deve ser evitada. Se a descrição de um abrir e fechar de asas numa ave é descrita com grande pormenor, com descrição das diferentes fases por que as asas passam, a inclinação do corpo, etc. e o comportamento de beber água é descrito apenas como «beber», dá origem a um desequilíbrio susceptível de distorcer os resultados. Neste sentido, as categorias devem ser homogêneas: o grau de particularização deve ser idêntico. Slater (1973) e Martin e Bateson (1986) descrevem os cuidados a ter na definição do catálogo de categorias comportamentais. Nos estudos que têm sido realizados há uma variação grande no tipo de categorias definidas. Em uns casos as categorias pertencem todas a uma mesma categoria funcional (cuidados individuais: Delius, 1969; Fentress & Stilwell, 1973). Noutros casos são incluídos todos os comportamentos (Jones & Brain, 1985; Stevenson & Poole, 1976). O problema levanta-se quando as categorias não são homogêneas entre si, havendo maior atenção sobre umas que outras, o que deve ser evitado. A questão do grau de particularização pode ser melhor entendida com o exemplo referido por Huntingford (1984) sobre a escolha de cores em pintos. Quando dão bicadas num fundo de determinada cor não dão apenas uma bicada, mas uma série delas. Assim será redundante contabilizar o número total de bicadas em cada cor. São as séries de bicadas que interessa medir porque reflectem a intenção de bicar.

Pode igualmente verificar-se que vários comportamentos ocorram simultaneamente o que torna difícil adaptar os dados a uma matriz de transições. No caso de os actos simultâneos serem muito frequentes, é preferível analisar os dados com técnicas de correlação (Slater, 1973).

Finalmente, as designações a atribuir às categorias não deverão ser conclusivas das consequências ou intenções do acto, mas ser o mais possível descritivas, para evitar introduzir juízos de valor *à priori* nas categorias; embora por vezes seja necessário utilizar designações funcionais para não perder informação essencial (Martin & Bateson, 1985).

Toda esta fase é bastante subjectiva e depende largamente da sensibilidade do investigador. Mesmo quando já existem reportórios de comportamento, ou etogramas descritos para a espécie que se estuda, é preferível construir um catálogo novo e analisar depois os pontos em que diverge dos de outros. Como foi referido atrás, a definição das categorias não é independente daquilo que se está a investigar.

#### FORMAS DE REGISTO DOS DADOS

O melhor processo de analisar sequências de comportamento, principalmente por causa de fases de transições muito rápidas, é utilizar um sistema de filmagem em vídeo que permite posterior análise pormenorizada. O registo das sequências gravadas em vídeo pode ser feito por transcrição oral para cassette, normalmente para o papel, ou directamente para um microcomputador através do teclado. Este processo é o que garante maior rapidez, mas exige uma longa habituação. Além disso está limitado a um reduzido número de categorias por causa da dificuldade em memorizar a posição das teclas. Martin e Bateson (1986) fazem uma excelente revisão dos diferentes métodos de registo.

Os registos podem ser dos actos de um só indivíduo, dos dois separadamente ou de

ambos de forma alternada, tudo dependendo do que se pretende fazer e do tratamento que se irá dar aos dados. Quanto a este aspecto dos registos, por exemplo Altmann (1965), que não pretendia analisar o papel desempenhado por cada indivíduo, fez uma análise de sequências de todo um grupo de macacos (*Macaca mulata*), usando um sistema de observação animalfocal. Nos casos mais simples de interações entre somente dois indivíduos, ou se opta por só registar os actos de um — o vencedor por exemplo — ou se registam os de ambos. No primeiro caso está presente a assumpção, raramente verdadeira, de que o oponente se comporta de uma forma bastante consistente com as acções do primeiro. Brain e colaboradores (Jones & Brain, 1985) solucionaram este problema utilizando, em interações agressivas em *Mus musculus*, linhagens não agressivas ou ratos anósmicos, que raramente mostram comportamentos agressivos, de forma a obter um oponente *standard* para os indivíduos cuja agressividade se pretendia estudar, o que é extremamente útil quando se pretende parametrizar os níveis e a intensidade da agressividade. Este sistema não é muito útil quando se pretende analisar a influência que os interactuantes exercem entre si, ou conhecer as características da própria interacção. Descrever o comportamento de ambos os interactuantes é a solução óbvia para esses casos, mas que não está liberta de problemas, como por exemplo a forma de integrar os dois registos (Slater, 1973).

#### ANÁLISE DE SEQUÊNCIAS DE COMPORTAMENTOS

##### 1. *Sequências de actos de um só indivíduo*

Separam-se aqui as sequências em que só os actos de um indivíduo são registados daquelas em que se registam os de dois ou mais porque o tipo de análise e os métodos a empregar são bastante diferentes. É evi-

dente que numa situação de interacção entre dois indivíduos podemos dispor de sequências individuais, bastando para isso registar separadamente os actos de cada um. Este procedimento tem vantagens por permitir a utilização de métodos mais eficazes além de ser extremamente difícil integrar os registos de ambos os indivíduos. Tem, no entanto, um problema: não permite estudar a influência que o oponente exerce, em cada momento, sobre o comportamento do primeiro indivíduo. A forma de resolver este problema será analisada na próxima secção.

Sendo os actos A, B e C as unidades de comportamento, pode começar-se por verificar quantas vezes A antecede B e C, quantas B antecede C e A, etc. Estes dados podem ser organizados numa matriz de transições como a representada na figura 1, sendo as linhas correspondentes ao primeiro acto e as colunas correspondentes ao segundo. Em cada célula é assinalado o número de vezes que cada acto antecede outro (o número superior). Estes dados podem ser analisados verificando se o número de associações entre eles é resultado do acaso ou tem uma base qualquer. A probabilidade de ocorrência de A associado a B, para um total de 1000 actos em que A acontece 100 vezes e B 10, é  $p = (10/1000) * (100/1000) = .001$ ; A-B ou B-A. Mas a probabilidade de A ocorrer após B, ou seja a probabilidade condicional de A dado B, é diferente:

$P(A_2 | B_1) =$  assume que há independência entre os eventos. Se

$P(A_2 | B_1) \neq P(A_2)$  então não há independência.

FIGURA 1

2º	1º	A	B	C	Total de linha
A		15 18.6	9 10.7	21 15.7	45
B		24 16.5	7 9.5	9 13.9	40
C		6 9.9	10 5.7	8 8.4	24
Total de coluna		45	26	38	109

Uma sequência de comportamentos pode ser descrita como sendo uma cadeia de Markov se as probabilidades de ocorrência de diferentes actos estiverem apenas dependentes do acto imediatamente precedente e não de quaisquer outros anteriores. Se

$$P(m | kl) = P(m | l)$$

então temos uma cadeia de Markov de primeira ordem, isto é, dada a sequência A-B-C, a probabilidade de C suceder a B não é de qualquer modo alterada pela natureza de A. Se a ocorrência de C depender, além de B, de r actos anteriores e não de outros temos markoviano de ordem r. A análise de dependências entre actos faz-se através da comparação com um modelo aleatório, verificando-se se algumas dessas transições são mais ou menos frequentes do que seria previsível se ocorressem ao acaso. Para tal calculam-se as frequências esperadas de cada transição (o número inferior em cada célula na figura 1) a partir da matriz de transições observadas:

$$E = \frac{\text{linha} * \text{coluna}}{\text{total geral}}$$

A análise dos desvios entre as transições esperadas e observadas (O - E) é feita com um teste de  $X^2$  de qualidade do ajustamento. Se as diferenças forem pequenas o  $X^2$  aproximar-se-á de zero. Se forem grandes isso deve-se à existência de dependências entre os actos. Para a matriz da figura 1  $X^2 = 13.34$ ,

$$X^2 = \sum_{i=1}^n \frac{(O-E)^2}{E}$$

sendo i cada uma das células da matriz. A comparação também pode ser feita através de teoria da informação (Chatfield & Lemon, 1970; Fentress & Stilwell, 1973; Steinberg, 1977) como veremos mais adiante. A realização de um teste de independência requer, no entanto, que poucas células se

apresentem vazias, não existam frequências esperadas inferiores a 1 e nunca mais de 20 % das células com valores inferiores a 5. Chatfield e Lemon (1970) propõem que se condensem algumas categorias com baixas ocorrências de forma a aumentar os valores das células. Jones e Brain (1985) optaram por eliminar os casos em que as transições eram inferiores a 5. Também se podem aumentar os valores de cada célula juntando as sequências de vários indivíduos na mesma matriz (Jones & Brain, 1985). Slater e Ollason (1972), no entanto, consideram que para estudos populacionais é preferível analisar cada sequência separadamente para que a variação individual não seja eliminada, o que nem sempre é possível — quando o catálogo de actos é vasto e as sequências curtas é necessário juntar os dados de vários indivíduos para aumentar os valores nas células da matriz. De um modo geral não é muito aconselhável misturar frequências obtidas a partir de sequências de diferentes comprimentos, embora no caso de sequências longas a diferença seja negligenciável.

Na matriz da figura 1 representam-se transições entre os próprios actos, ou seja A pode suceder a A — mas é muito difícil determinar se um acto terminou e se reiniciou ou se continua em curso, o que resulta numa situação de extrema subjectividade apenas devida ao observador e ao critério que optou por empregar.

FIGURA 2

2°	1°	A	B	C	Total de linha
A		100 87.1	9 17.4	21 25.5	130
B		24 26.8	7 5.4	9 7.8	40
C		6 16.1	10 3.2	8 4.7	24
Total de coluna		130	26	38	194

Uma mera mudança na matriz passando A-A de 15 para 100 (figura 2) implica uma alteração completa no cálculo das probabilidades esperadas em todas as células, pelo que é preferível não considerar a possibilidade de um comportamento se suceder a si mesmo, deixando a diagonal em branco. Também as transições impossíveis — se um animal tem que se deslocar para passar do bebedor ao comedor não pode haver transições entre beber e comer — devem ser assinaladas e consideradas vazias. Essas células não devem ser consideradas para o cálculo das probabilidades esperadas, porque alterariam os valores calculados para as outras células (Slater & Ollason, 1972). Lemon e Chatfield (1971) descrevem um processo para calcular as frequências esperadas sem considerar as células em branco.

Verificando-se existirem dependências entre os actos, torna-se necessário testar se uma cadeia de Markov de 1.ª ordem serve para descrever as frequências ou seja

$$P(A_3 | B_1 C_2) = P(A_3 | C_2)$$

No caso de

$$P(A_3 | B_1 C_2) \neq P(A_3 | C_2)$$

então B também exerce influência em A pelo que deve haver uma ordem de dependência mais elevada. O teste é também feito com o  $X^2$  de qualidade do ajustamento. O mesmo procedimento pode ser utilizado para estudar dependências de ordem superior. Mas a utilização do  $X^2$  torna-se mais problemática à medida que a ordem aumenta. O número de sub-sequências possíveis com comprimento p é dado por  $C^p$ . Assim, para  $C=10$  há 1000 tripletos ( $10^3$ ) possíveis e muitos terão frequência nula mesmo quando a sequência é demasiado longa, o que impede o uso do  $X^2$ . No entanto, se o número de actos for pequeno e houver bastantes dados, o teste pode ser usado até um processo de 3.ª ordem (Chatfield & Lemon, 1970). Este procedimento não fornece mais do que uma imagem geral do padrão de dependências entre actos e está limitado por várias condições que, como veremos, rara-

mente são possíveis em estudos de comportamento.

Para uma análise mais pormenorizada das transições é relevante detectar quais as transições que são significativamente mais comuns do que os seus valores esperados. Com este procedimento é possível obter dados indicativos sobre as relações entre os diferentes comportamentos e esta informação é mais interessante, em termos comportamentais, do que a fornecida pela análise de toda a matriz de transições, sobre a existência ou ausência de algum tipo de dependência entre os actos de um modo geral. O processo consiste no cálculo de um  $X^2$  sobre a matriz condensada numa tabela de 2\*2 para a célula considerada, verificando-se se a transição em questão é mais frequente do que o esperado. Pode também analisar-se a matriz de transições detectando assimetrias em certas sequências. Se os actos se sucedem ao acaso, espera-se que as frequências de transição do tipo A-B e B-A sejam iguais. Os casos em que isso não ocorre podem significar a existência de dependências. As diferenças são testáveis com o teste de «pares cruzados» de Wilcoxon (Slater & Ollason, 1972).

#### O REQUISITO DE ESTACIONARIDADE

Há um factor que limita fortemente a aplicação da análise de Markov a dados comportamentais que é o requisito de estacionaridade. Para que uma sequência seja estacionária é necessário que as probabilidades de ocorrência de cada acto não mudem ao longo do tempo, isto é, têm que ser as mesmas em qualquer ponto de sequência. Em termos de comportamento, seria considerar que o animal se encontra num estado estacionário, o que é altamente improvável (Bekoff, 1977; Oden, 1977). Seria presumir, por exemplo, uma completa ausência de motivação no animal. Qualquer tendência é um sinal de não estacionaridade. É muito

mais acertado supor que, quer o vencedor final quer o vencido de uma interacção agressiva, alteram as probabilidades com que executam certos actos enquanto o combate se desenrola que o contrário.

A ausência de estacionaridade pode conduzir a matrizes de transição muito complexas: se as probabilidades de transição não são as mesmas ao longo de toda a sequência, a matriz será a soma de vários processos que se confundirão entre si.

Slater (1973) e Bekoff (1977) expõem as principais consequências de se comparar os dados desta forma sem obedecer ao requisito de estacionaridade. Como o primeiro refere, quanto menos válidas as assumpções mais irreais serão os resultados.

#### ALTERNATIVAS

É possível, de certo modo, solucionar o problema dividindo a sequência em várias sequências de menor comprimento e analisá-las separadamente, como Hazlett e Estabrook (1974) fizeram com os combates de uma espécie de caranguejos, o que garante menor variação nas probabilidades de ocorrência. No entanto isso requer uma quantidade muito maior de dados por causa do problema das células com frequências muito baixas.

Alguns procedimentos alternativos foram já referidos atrás. Bekoff (1977), por seu lado propõe a utilização do Coeficiente de Variação (CV) —  $CV = \text{Desvio Padrão} * 100 / \text{Média}$  — para analisar a variabilidade existente numa sequência e comparar essa variação em sequências em condições distintas. Utilizou o CV num estudo comparativo do comportamento social de jovens coiotes tendo verificado serem as sequências em situação de jogo mais variáveis por comparação com interacções agonísticas.

Outro modo interessante de rodear o problema da estacionaridade foi desenvolvido por Jones e Brain (1985) e baseia-se no cál-

culo de um índice ou coeficiente de distância entre elementos a partir de transições. Os coeficientes de distância são utilizados para realizar uma análise de aglomeração ('cluster analysis') entre elementos. Deste modo é possível analisar a semelhança posicional entre os actos, ou seja, aqueles que se assemelham em termos de posição relativa da sequência e não necessariamente aqueles que se seguem. A figura 3 mostra a comparação entre cada par de elementos da matriz, na sua relação com os outros, através do cálculo de um Qui-quadrado. O valor de  $X^2$  global é tomado como medida de similaridade entre os elementos. Aliás, como o valor é tanto mais alto quanto menor a semelhança entre os elementos, este funciona como medida de distância inter-elementos (Sneath & Sokal, 1973). As medidas de distância assim obtidas são dispostas numa matriz de distâncias, simétrica, que serve para a análise de aglomerações. O resultado é um dendograma como o representado na figura 4, em que se estabelecem graus de associação, ou similaridade, entre os comportamentos, possibilitando, por exemplo, o estabelecimento de famílias de comportamentos que podem ter uma base funcional importante. Esses resultados são mais úteis na formulação de novas hipóteses para prosseguir a investiga-

ção do que no estabelecimento de conclusões.

FIGURA 3

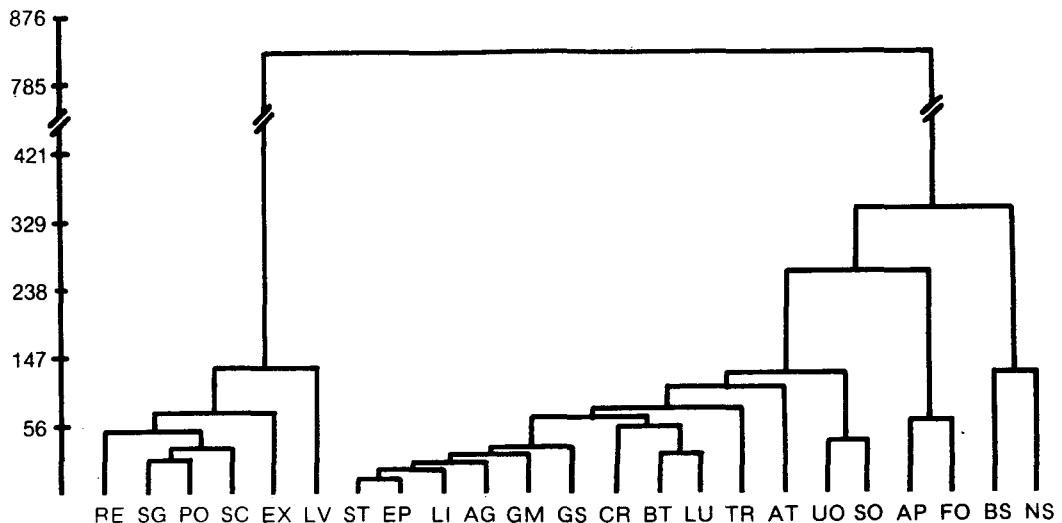
	A/B	C	D	Total de linha
A	15 9.5	7 8.0	6 10.5	28
B	4 9.5	9 8.0	15 10.5	28
Total de coluna	19	16	21	Grande total 56

$$X^2_{A,B} = 10.48$$

Outros métodos para detectar associações entre elementos, embora não directamente relacionados com a sequenciação, têm igualmente sido empregues. Referem-se aqui dois deles: análise factorial e correlação.

A análise factorial extrai do conjunto de variáveis um número reduzido de factores (3 ou 4) os quais são susceptíveis de explicar a maior parte da variação existente nas variáveis originais. As variáveis têm que ser independentes, sendo as correlações entre elas explicadas através desses factores. A análise factorial, ao contrário da análise de Markov, não dá importância à sequência particular

FIGURA 4



dos actos, procurando relacioná-los por aquilo que podemos designar por «estado motivacional» e que lhes está subjacente (Slater, 1973). A associação entre os actos é entendida como resultado da existência de factores causais comuns subjacentes aos comportamentos. Aspey e Blankenship (1977) usaram análise factorial para interpretar o comportamento agonístico numa espécie de aranhas. Da análise extraíram 4 factores que explicavam 74,3 % da variância existente. Em função dos comportamentos correlacionados com cada factor procuraram atribuir-lhes um significado biológico tangível: aproximação/sinalização, perseguição vigorosa, correr/fugir. Embora neste caso se obtenha uma imagem relativamente clara da estrutura motivacional durante a interacção, nem sempre assim é. Não é fácil atribuir um significado a esses factores complexos. Como Slater (1973) afirma «é duvidoso que a extracção de factores que são em si mesmos de causalidade complexa permita avançar algum conhecimento».

As técnicas de correlação constituem outro modo de analisar associações entre elementos, mas só podem ser aplicadas a sequências temporais. Os métodos até agora descritos não requeriam que as sequências fossem medidas numa escala temporal. A correlação entre cada par de actos é calculada em função das suas frequências numa série de unidades temporais, determinando-se se estão positiva ou negativamente associados. A estacionaridade é aqui um problema menor (Slater, 1973). A principal dificuldade reside na definição da unidade de tempo a ser considerada, porque os resultados dependem totalmente disso. A solução é calcular as frequências considerando várias unidades de tempo e verificar depois qual a mais adequada. Uma análise deste tipo dá informação apenas acerca da tendência para certos comportamentos ocorrerem associados, mas não fornece qualquer indicação sobre a ordem pela qual tendem a seguir-se.

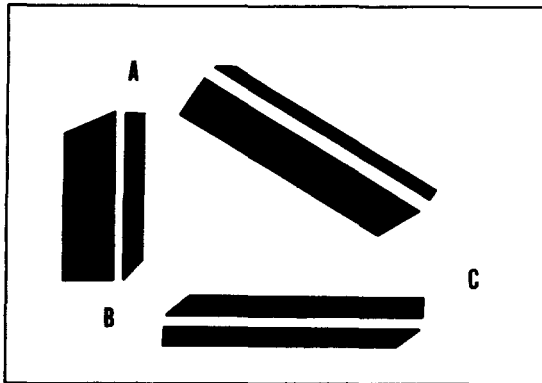
Para isso torna-se necessário proceder a uma análise de intercorrelações entre os elementos.

O cálculo de intercorrelações e de autocorrelações difere dos coeficientes de correlação normais porque se consideram vários intervalos de tempo sucessivos e em ordem crescente (Delius, 1969). As correlações são calculadas entre dois pontos após um intervalo que varia de forma sistemática. O espectro que se obtém poderá corresponder a uma curva sinusoidal se houver um padrão cíclico de repetição dos comportamentos. Se um comportamento tender a suceder ciclicamente outro com um intervalo de 8 seg. e o intervalo mínimo for de 1 seg., quando o intervalo atingir a diferença de 8 segundos surgirá uma correlação positiva. Delius (1969) pôde encontrar vários ritmos em diferentes comportamentos relacionados com cuidados corporais utilizando ambos os métodos em *Alauda arvensis*. Enquanto a autocorrelação é simétrica para intervalos positivos e negativos, a intercorrelação tem que ser calculada para ambos os tipos de intervalo. As técnicas de correlação, como se disse atrás, são particularmente importantes quando se averiguam mecanismos causais. Não é eficiente na interpretação da sequência dos comportamentos quando esta não depende tanto de um mecanismo causal subjacente.

Há um conjunto de procedimentos mais simples, cujas exigências de aplicabilidade são mínimas, mas se revelam extremamente úteis. A construção de diagramas de fluxo entre comportamentos (Ewing, 1975; Slater & Ollason, 1972) permite verificar quais as transições mais frequentes e as relações que se estabelecem entre os actos. Slater e Ollason (1972), a partir de matrizes de transições, construíram diagramas de fluxo para as transições significativamente mais frequentes do que o esperado (teste binomial), conseguindo uma melhor visualização das relações entre os diferentes comportamentos. Também Aspey e Blankenship (1977) procuraram

maior percepção das relações entre os comportamentos, calculadas com análise factorial, utilizando diagramas de fluxo.

FIGURA 5



Na figura 5 representam-se os valores das frequências observadas da matriz da figura 1, sem as transições entre os próprios actos. A espessura das setas indica a intensidade da transição. No estudo de Poole e Fish (1975), sobre comportamento de jogo em *Rattus norvegicus*, os diagramas foram construídos a partir dos valores do desvio normal calculado com o teste binomial.

O cálculo das frequências acumuladas de cada comportamento e a sua representação num sistema de eixos fornece informação preliminar bastante interessante. Com um grau variável de comportamento para comportamento, é possível detectar a existência de regiões de maior ou menor ocorrência de certos comportamentos, numa interacção agressiva por exemplo. A frequência com que certos actos ocorrem pode ter uma variação brusca em certos pontos da sequência (Figura 6), o que pode corresponder a pontos de decisão da sequência, de redução da incerteza (Dawkins & Dawkins, 1973). Esses pontos podem, inclusivamente, ser utilizados para subdividir a sequência, de modo a reduzir a não estacionaridade e a possibilitar a análise com cadeias de Markov, como Hazlett e Estabrook (1974) realizaram.

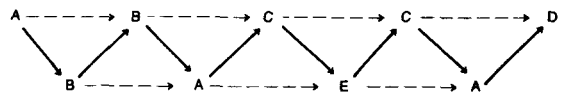
FIGURA 6



### 1. Sequências de interacções entre indivíduos

Muitos dos estudos sobre sequências de interacções sociais têm como objectivo demonstrar que o comportamento de um animal é afectado pelo de outros e obter assim evidência sobre a existência de comunicação. Parecerá à primeira vista que isso é fácil de conseguir com o estabelecimento de uma matriz de transições entre dois indivíduos. No entanto o problema da estacionaridade numa matriz desse tipo é ainda mais grave, porque um dos principais aspectos da comunicação é precisamente a alteração das probabilidades de ocorrência dos comportamentos (Slater, 1973).

FIGURA 7



Como o gráfico da figura 7 mostra, o comportamento é sempre influenciado por factores externos e internos. Os seus actos podem depender dos do oponente, mas não são independentes do que o próprio fez anteriormente.

A importância de cada um dos factores na determinação das decisões de cada indivíduo pode ser percebida com a utilização da teoria de informação. Consegue-se calculando

separadamente medidas de transferência de informação para os dois tipos de transição da figura 6 (A-B-C-C-D; A-B-B-A-C-E-A-D), o que possibilita o estudo da comunicação inter e intra-individual. É possível testar hipóteses acerca da existência de comunicação entre animais e quantificar o valor dessa informação. Quando um animal faz certos movimentos ou emite sons, somos facilmente levados a pensar que está a comunicar. Mas, a menos que consigamos provar que estes actos alteram a distribuição de probabilidades dos actos dos outros indivíduos, não podemos garantir que ela exista.

A unidade de informação ou 'bit' equivale à quantidade de informação necessária para escolher entre duas alternativas de igual probabilidade. Havendo 4 alternativas igualmente prováveis a escolha requer 2 bits de informação e com  $n$  alternativas seriam necessários  $\log_2 n$  bits. No entanto, os actos não são escolhidos com igual probabilidade pelos animais, a partir do seu repertório comportamental, pelo que a probabilidade de ocorrência de um acto deve ser considerada no cálculo da quantidade de informação. O conteúdo de informação de uma série (X) de observações de comportamento (ou seja a dificuldade em prever essa sequência) é calculada com

$$H(x) = - \sum_{i=1}^n p(i) \log_2 p(i)$$

em que  $i$  representa os actos do repertório e  $p(i)$  é a probabilidade do acto  $i$ . Como raramente essa é conhecida usa-se o seu estimador ( $\hat{p}$ ) dado pelo  $n^\circ$  de ocorrências de  $i$ /total  $n^\circ$  de ocorrências (Steinberg, 1977). É possível, então, usar estes conceitos para medir a comunicação entre animais. Tendo um registo de actos consistindo numa série de acções de um animal A e uma série de actos consequentes de B, o conteúdo de informação da série de cada um individualmente obtém-se a partir da fórmula referida acima, de que resulta  $H(A)$  e  $H(B)$ . A partir

da matriz de transições pode calcular-se a informação contida no conjunto de pares de sucessivos actos ( $H(A,B)$ ) por um processo análogo. A incerteza condicional, ou seja a informação presente na distribuição de actos praticados por B, pode calcular-se com a fórmula

$$H(B | A) = - \sum_{ij} p(i,j) \log_2 p(j | i)$$

em que  $p(i,j)$  é a probabilidade incondicional de  $i$  e  $j$  associados e  $p(j | i)$  é a probabilidade condicional de  $j$  dado  $i$  (Huntingford, 1984).  $H(B | A)$  mede a quantidade média de incerteza acerca do comportamento do animal B que permanece quando as acções precedentes de A são conhecidas. Se o comportamento de B for totalmente previsível a partir do conhecimento dos de A, então  $H(B | A)$  será igual a 0. A diferença entre  $H(B)$  e  $H(B | A)$  designa-se transmissão ( $T(A,B)$ ) e mede a informação transmitida de A para B.

Chatfield e Lemon (1970) e Steinberg (1977) apresentam com clareza a aplicação da teoria de informação ao comportamento animal e o segundo faz uma extensão do método para interacções entre vários indivíduos. Huntingford (1984) descreve uma série de aplicações a estudos etológicos.

A utilização da teoria de informação não está livre de problemas sendo o principal, embora solucionável (Steinberg, 1977; Huntingford, 1984), a exigência de estacionaridade. Oden (1977) desenvolveu um método de análise das dependências entre actos, em interacções de dois indivíduos, assumindo à partida a sua não-estacionaridade. O procedimento — que não se expõe aqui dada a sua complexidade — permite calcular graus de dependência entre actos, como se de cadeias de Markov se tratasse, e transformá-los ainda em medidas de informação. A sua aplicabilidade restringe-se somente a sequências curtas e exige grande quantidade de dados.

Nenhum modelo é perfeito para solucionar todos os problemas da análise de seqüências e obter a informação pretendida dos dados. Muito do que se pode conseguir com os métodos descritos depende do reconhecimento das restrições de cada um e da forma como os dados são obtidos, pelo que é extremamente importante saber a que tipo de tratamento estatístico se vai recorrer na fase de delineamento da experiência.

### Agradecimentos

Agradeço aos meus colegas Dr. Vitor Almada e Dr. Luís Vicente bem como ao Dr. George Estabrook as proveitosas discussões que tivemos e que foram fundamentais para a elaboração deste material. Os meus agradecimentos também ao ISPA e à Sociedade Portuguesa de Etologia pela oportunidade concedida de expôr estas considerações.

### BIBLIOGRAFIA

- ALTMANN, S. A. (1965). Sociobiology of rhesus monkeys. II. Stochastics of social communication. *J. Theoret. Biol.* 8: 490-522.
- ASPEY, W. P. & BLANKENSHIP, J. E. (1977). Spiders & snails & statistical tales: Application of multivariate analyses to diverse ethological data. In: B. Hazlett (ed.) *Quantitative methods in the study of animal behaviour*. New York: Academic Press.
- BEKOFF, M. (1977). Quantitative studies of three areas of classical ethology: Social dominance, behavioral taxonomy, and behavioral variability. In: B. Hazlett (ed.) *Quantitative methods in the study of animal behavior*. pp. 1-46. New York: Academic Press.
- CHATFIELD, C. & LEMON, R. E. (1970). Analysing sequences of behavioural events. *J. Theoret. Biol.* 29: 427-445.
- DAWKINS, M. S. (1983). The organisation of motor patterns, in T. R. Halliday & P. J. B. Slater (eds.) *Animal Behaviour*, vol. 1, Causes and Effects, pp. 75-99. Oxford: Blackwell Scientific Publications.
- DAWKINS, R. & DAWKINS, M. (1973). Decisions and the uncertainty of Behaviour. *Behaviour*, 45: 83-103.
- DELIUS, J. D. (1969). A stochastic analysis of the maintenance behaviour of skylarks. *Behaviour* 33: 137-178.
- EWING, A. W. (1975). Studies on the behaviour of cyprinodont fish. II the evolution of aggressive behaviour in old world rivulins. *Behaviour*, 52: 172-195.
- FENTRESS, J. C. & STILWELL, F. P. (1973). Grammar of a movement sequence in inbred mice. *Nature*, vol. 244, July, pp. 52-53.
- HAZLETT, B. A. & ESTABROOK, G. (1974). Examination of agonistic behaviour by character analysis II. Hermit crabs» *Behaviour* 49: 88-100.
- HUNTINGFORD, F. A. (1984). *The study of Animal Behaviour*, London: Chapman and Hall.
- JONES, S. E. & BRAIN, P. F. (1985). An illustration of simple sequence analysis with reference to the agonistic behaviour of four strains of laboratory mouse. *Behavioural Processes*, 11, pp. 365-388.
- LEMON, R. E. & CHATFIELD, C. (1971). Organization of song in cardinals. *Animal Behaviour* 19: 1-17.
- MARTIN, P. & BATESON, P. (1986). *Measuring behaviour an introductory guide*. Cambridge: Cambridge University Press.
- ODEN, N. (1977) — Partitioning dependence in nonstationary behavioral sequences; in B. Hazlett (ed.) *Quantitative methods in the study of animal behavior*, New York: Academic Press.
- POOLE, T. B. & FISH, J. (1975). An investigation of playful behaviour, in *Rattus norvegicus* and *Mus musculus* (Mammalia) *J. Zool.*, London, 175, pp. 61-71.
- SIMPSON, M. J. A. (1969). The display of the siamese fighting fish, *Betta splendens*. *Anim. Behav. Mono* 1: 1-73.
- SLATER, P. J. B. & OLLASON, J. C. (1972). The temporal pattern of behaviour in isolated male zebra finches: Transition analysis, *Behaviour* 42: 248-269.
- SLATER, P. J. B. (1973). Describing sequences of behaviour, in Bateson & Klopfer (eds.) *Perspectives in ethology*. Plenum Press, N. Y., London, 131-153.
- SNEATH, P., SOKAL, R. R. (1973). *Numerical taxonomy*. S. Francisco: W. H. Freeman & Co.
- STEINBERG, J. B. (1977). Information theory as an ethological tool, in B. Hazlett (ed.) *Quantitative methods in the study of animal behavior*, Academic Press, New York, London, pp. 47-74.
- STEVENSON, M. F., POOLE, T. B. (1976). An ethogram of the common marmoset (*Calitrix jacchus jacchus*: general behaviour repertoire. *Animal Behaviour*, 24: 428-451.