

2. REFERENCIAL TEÓRICO

O objetivo desta seção é o de apresentar e revisar os conceitos da literatura considerados necessários ao estudo do relacionamento dos recursos das empresas e o seu posicionamento estratégico na arena competitiva, como resultado da exploração de seus recursos, intencionalmente ou não. A caracterização do ambiente competitivo, e sua influência no desenvolvimento de vantagem competitiva sustentável; a Visão *Resource-Based*, de seus fundamentos à contribuição da formulação da estratégia das empresas; o conceito “grupo estratégico”, enfatizando a inimitabilidade e imobilidade dos recursos como os fatores de caracterização destes grupos, são os tópicos abordados.

Serão também apresentadas e discutidas, de forma breve e introdutória, as técnicas de análise de clusters e de inteligência computacional, e sua adequação à utilização no estabelecimento de grupos estratégicos. O objetivo principal destes tópicos é permitir a comparação entre o modelo neuro-fuzzy a ser utilizado, e os modelos mais consagrados da análise de clusters.

2.1. CONCEITOS CENTRAIS DE ESTRATÉGIA

2.1.1. Ambiente Competitivo

A evolução das pesquisas sobre a estratégia das empresas, considerando as inúmeras definições, proposições e conceitos existentes na literatura sobre o que isto signifique, apresentam como base à questão da competição, e como obter sucesso, neste ambiente de competição. Obter sucesso na competição significa a superação dos competidores que conseguem apenas a equivalência e a desvantagem competitiva, ou seja, obter vantagem competitiva (BARNEY, 2002, p. 25), e levando-se em consideração a dinâmica e complexidade do ambiente competitivo, ter a capacidade de sustentá-la e renová-la (DAY; REIBSTEIN, 1999, p. 59). Esta superação é evidenciada através dos indicadores de performance, também objetos de inúmeras pesquisas no campo da administração estratégica, mas que de forma geral estão relacionados aos valores econômicos que as empresas criam pela aplicação de seus ativos produtivos. Questões que surgem naturalmente, entre outras, são: quais são os indicadores de performance mais

adequados? Quais são os fatores que influenciam na performance da empresa: os organizacionais ou os ambientais? Não são questões fáceis de responder e que com certeza não admitem uma solução única e definitiva.

A natureza dos indicadores – se financeiros, operacionais ou ambos, e ainda, se quantitativos ou qualitativos – irá delimitar o domínio de avaliação da performance, e conseqüentemente, o domínio onde estarão sendo confrontados os resultados das estratégias das empresas. Para que a avaliação seja a mais abrangente possível, é importante que outros aspectos, além dos financeiros e operacionais, sejam considerados (VENKANTRAMAN; RAMANUJAM, 1986). Uma abordagem está em considerar duas dimensões dos indicadores de performance: a abrangência de seu domínio, limitado à financeira e a operacional, e a natureza da fonte de informação, se primária ou secundária. Essas duas dimensões permitem obter dez formas diferentes de tratar a questão da performance: (1) indicadores financeiros utilizando fontes primárias, (2) indicadores financeiros utilizando fontes secundárias, (3) indicadores operacionais utilizando fontes primárias, (4) indicadores operacionais utilizando fontes secundárias, (5) indicadores financeiros utilizando fontes primárias e secundárias, (6) indicadores operacionais utilizando fontes primárias e secundárias, (7) indicadores financeiros e operacionais utilizando fontes primárias, (8) indicadores financeiros e operacionais utilizando fontes secundárias, (9) indicadores financeiros utilizando fontes primárias e indicadores operacionais utilizando fontes secundárias, e (10) indicadores financeiros utilizando fontes secundárias e indicadores operacionais utilizando fontes primárias. Cada um destes tipos apresenta vantagens e desvantagens, bem como uma metodologia de aplicação, devendo a sua utilização ser adequada à abrangência desejada nos estudos e a disponibilidade dos dados. O compromisso parcimônia x poder explanatório é um aspecto a ser observado no caso da busca por uma maior abrangência de domínio (VENKANTRAMAN; RAMANUJAM, 1986).

A questão dos fatores de influência é decorrente da natureza das pesquisas realizadas e teorias econômicas, como a economia das organizações industriais, que resultaram, e ainda resultam, em indicadores predominantemente quantitativos

(BARNEY, 2002, p. 28-68), orientando o desenvolvimento de teorias e abordagens com a ênfase no ambiente externo e seus agentes (MINTZBERG; QUINN, 1996, p.47-55, p. 75-83) para a formação e avaliação da adequação das estratégias das empresas. Duas considerações são importantes nas abordagens que enfatizam o ambiente. Primeiro, os recursos das empresas não são ignorados, mas colocados a serviço da análise *SWOT*, a avaliação dos pontos fortes e dos pontos fracos da organização a luz das oportunidades e ameaças em seu ambiente, e também trazendo as primeiras referências à importância dos recursos e competências distintas. Segundo, é que tais conceitos quando aplicados a ambientes competitivos relativamente estáticos e pouco complexos permitem avaliações adequadas, tendo sido amplamente utilizados e ainda de grande valor para o planejamento estratégico das empresas.

A evolução da competição em direção à complexidade e à dinâmica das arenas competitivas refletiu-se nas pesquisas de gerenciamento estratégico, no sentido de buscar abordagens alternativas para a obtenção da vantagem estratégica. A visão de que recursos e produtos são as faces de uma mesma estrutura é um dos pontos de partida destas pesquisas. Ao estabelecer o nível de atividades de uma empresa em mercados de diferentes produtos, é possível inferir o comprometimento mínimo de recursos, assim como, ao estabelecer um perfil de utilização de recursos, é possível buscar um mercado ótimo para os produtos. Sua correlação com a performance está associada a responder a questão: Sob que circunstância poderá um recurso trazer altos retornos por longos períodos? (WERNERFELT, 1984). Outro aspecto interessante é a necessidade de avaliação de todos os recursos, tangíveis ou não. A importância da apreensão, acumulação e utilização destes recursos – principalmente dos intangíveis – na obtenção da vantagem competitiva foram abordadas por (WERNERFELT, 1989), (COOL; DIERICKX, 1989), (PRAHALAD; HAMEL, 1990), (HALL, 1993) e (BARNEY, 2002, p.160-171).

Pesquisas empíricas também buscaram verificar os fatores importantes na obtenção de uma performance acima do normal, não obtendo evidências fortes da predominância de fatores ambientais, representados pelos fatores econômicos, ou

dos fatores organizacionais. Entretanto, conclusões como: os fatores organizacionais permitem uma melhor explicação para a variância da performance do que os econômicos (HANSEN; WERNERFELT, 1989), (HAWAWINI; SUBRAMANIAN; VERNDIN, 2003) e de que os fatores intra-indústria, os ambientais, afetam de forma diferente as empresas, isto é, enquanto os fatores organizacionais têm maior impacto nas empresas líderes e nas de pior desempenho, os fatores ambientais afetam as empresas com performance normal (HAWAWINI; SUBRAMANIAN; VERNDIN, 2003). Adicionalmente, um modelo que conjugou os dois fatores, econômico e organizacional, apresenta performance superior, indicando características suplementares entre os dois fatores (HANSEN; WERNERFELT, 1989).

Desta forma, em um ambiente competitivo complexo e altamente dinâmico, a abordagem “Como direcionar os recursos acumulados da empresa para competir neste ambiente?“, poderá não lograr êxito na busca por uma performance acima do normal, uma vez que todos os competidores estarão buscando o mesmo objetivo. Enquanto que a abordagem “Com os recursos acumulados, qual será a melhor arena competitiva para atuar e então a partir daí, como competir?“, poderá levar a uma performance acima do normal, como resultado de uma postura da empresa no sentido de buscar a vantagem competitiva sustentável.

2.1.2. A Visão *Resource-Based* – Os Recursos da Organização e o Gerenciamento de sua Estratégia

Esta seção revisa os conceitos da Visão *Resource-Based* (WERNERFELT, 1984) através da seguinte abordagem:(1) apresentação e discussão de seus conceitos básicos, (2) a correlação dos recursos com a vantagem competitiva sustentável e (3) suas implicações no gerenciamento estratégico das empresas.

2.1.2.1. A Visão *Resource-Based*

O conceito *Resource-Based* foi inicialmente proposto por Berger Wernerfelt (WERNERFELT, 1984) como resultado de uma abordagem estratégica

das idéias de Penrose, sobre o crescimento das empresas (Penrose, 1959), na qual uma visão mais abrangente do que sejam considerados recursos produtivos é defendida, pela inclusão de características menos tangíveis das empresas, suas implicações competitivas e da heterogeneidade das empresas, ainda que na mesma indústria (BARNEY, 2002, p.154-155). Neste trabalho seminal, Wernerfelt propõe que: (1) as empresas sejam avaliadas pela perspectiva de seus recursos produtivos, e não de seus produtos, de forma a identificar aqueles que conduzam a uma maior lucratividade; (2) que as empresas que possuem estes recursos, e aproveitam o seu potencial, impõem uma barreira aos outros concorrentes de forma análoga às barreiras de entrada; (3) que a estratégia das grandes empresas deve conduzir a um equilíbrio entre a exploração dos recursos atuais e o desenvolvimento de novos, para a manutenção de uma melhor performance; (4) que uma aquisição pode ser vista como a compra de recursos em um mercado altamente imperfeito, aumentando a oportunidade à aquisição de recursos que permitam altos retornos. O resultado foi uma abordagem para a estratégia de crescimento dos recursos produzidos, simultaneamente com os produtos, porém indicando um caminho para a pesquisa dos outros recursos da empresa, bem como ressaltando a necessidade de se pesquisar sobre as dificuldades práticas de se identificar os recursos produtivos que conduzem a uma maior lucratividade.

Com relação à identificação dos recursos mais importantes, destacam-se três abordagens, todas contextualizadas na questão da formulação de uma estratégia corporativa e na vantagem competitiva. Uma foi apresentada por Wernerfelt (WERNERFELT, 1989), sendo compostas de duas fases. Na primeira, os recursos críticos devem ser identificados e na segunda classificados quanto à sua capacidade. A forma de realizar a primeira fase é através da análise da resposta a três questões básicas: Quais recursos da empresa são únicos? Alguma equipe na empresa apresenta remuneração inadequada com o seu desempenho? Existe algum fornecedor, ou cliente, responsável pelo fornecimento, ou consumo, da maioria dos recursos? As respostas poderão não atender às expectativas dos gerentes, tornando sua análise mais complexa. Para classificar os recursos críticos de acordo com a “quantidade” disponível, são propostos três grupos: ativos fixos como, instalações fixas e equipamentos, empregados especializados em processos específicos e etc,

considerados disponíveis por longo prazo; os recursos conhecidos por *blueprints* como patentes, marcas e reputação, que apresentam disponibilidade praticamente ilimitada e culturas, relacionados ao resultado apresentados pelas equipes de trabalho, com disponibilidade limitada no curto prazo, isto é, decorrente dos resultados tangíveis de um projeto realizado por uma equipe multidisciplinar, por exemplo, porém ilimitada no longo prazo, decorrente da interação entre os membros da equipe. O aspecto interessante desta classificação está no fato de recursos tangíveis e intangíveis pertencerem a um mesmo grupo, sendo certamente sua aplicabilidade dependente da indústria e do ambiente competitivo.

Outra abordagem, desenvolvida por Barney e conhecida como VRIO, *Valuable, Rare, Inimitable e Organization* (BARNEY, 2002, p.158-185), tem como objetivo facilitar a identificação dos recursos e capacidades, avaliar sua heterogeneidade e imobilidade para associar as forças e fraquezas das empresas. Os critérios de escolha são o valor e a raridade dos recursos no ambiente competitivo, a dificuldade em ser copiados e a estrutura da empresa para aproveitar estas características. Para a condução da análise dos recursos são propostas as seguintes questões: (1) a questão do valor: os recursos e capacidades da empresa permitem-na responder às ameaças e oportunidades do ambiente competitivo? (2) a questão da raridade: os recursos de valor são controlados por poucos concorrentes? (3) a questão da inimitabilidade: as empresas que não possuem estes recursos enfrentarão desvantagens de custo para copiá-los ou desenvolvê-los? (4) a questão da organização: as políticas e processos da empresa são adequados para garantir a exploração dos recursos de valor, raros e difíceis de copiar?

A terceira, proposta por Grant (GRANT, 1991), tem na escolha dos recursos o primeiro passo para a formulação da estratégia da empresa. Propõe, como forma de facilitar a análise, que os recursos sejam classificados em seis categorias: os físicos, os humanos, os financeiros, os tecnológicos, a reputação e os organizacionais. Estes recursos devem então ter suas forças e fraquezas comparadas com os recursos dos outros competidores, e identificadas oportunidades em que sua utilização levará à melhor rentabilidade. Nesta abordagem, as empresas também devem ser capazes de transformar estes recursos

em capacidades distintivas, entendo-se por capacidade não apenas o agrupamento de recursos e pessoas, mas também uma complexa tarefa de gerenciar seus inter-relacionamentos.

As abordagens propostas para a identificação dos recursos evidenciam uma questão central da Visão *Resource-Based*: os recursos que conferem às empresas alguma vantagem competitiva não devem ser facilmente identificados pelo ambiente competitivo. Quando identificados, devem apresentar dificuldades, preferencialmente impossibilidade, em sua reprodução pelos competidores ou mobilidade imperfeita, entendida como a peculiaridade de não apresentar o mesmo valor para as outras organizações, para o ambiente competitivo, isto é, apresentar ambigüidade causal na sua aplicação. Por ambigüidade causal entende-se a ambigüidade intrínseca aos resultados, ações e recursos, e que impossibilita a obtenção de uma correlação precisa entre os mesmos.

Na busca para obter um tratamento analítico para a ambigüidade causal, dois trabalhos merecem destaque por apresentarem conceitos que servem de base para esta pesquisa. No primeiro, a ambigüidade atua como um bloqueio à imitação e à mobilidade de fatores de produção, sustentando não ser possível obter uma lista finita, e sem ambigüidade, de fatores de produção responsáveis pelo sucesso de algumas empresas. De uma forma geral: os fatores de produção só podem ser “movimentados” se forem conhecidos (LIPPMAN; RUMELT, 1982). No outro, a ambigüidade causal é resultado do conhecimento tácito, da complexidade do inter-relacionamento entre recursos, competências e especificidade de ativos da organização. A ambigüidade causal é responsável pelo estabelecimento de barreiras à imitação, sendo sua “altura” e persistência determinadas pela agressividade do ambiente competitivo e pelo grau de ambigüidade (REED; FILLIPI, 1990).

A conclusão dos conceitos apresentados é que a simples posse, ou controle, de recursos – principalmente àqueles que apresentem valor para o ambiente competitivo – não garante a vantagem competitiva. Tais recursos deverão ter seus potenciais explorados de tal forma que não fique claro como o “valor” dos mesmos

é extraído, sob pena de serem copiados e transportados para o ambiente competitivo. Reside aí o conceito chave da Visão *Resource-Based*, isto é, o gerenciamento adequado dos recursos da empresa de forma a preservar sua inimitabilidade e imobilidade para o ambiente competitivo.

2.1.2.2. A Visão *Resource-Based* e a Vantagem Competitiva Sustentável

Discutidas as características, inimitabilidade e imobilidade, que devem os recursos da organização apresentar para merecer apreciação, da perspectiva do planejamento estratégico, cabe então a questão: a exploração dos recursos de maior grau de inimitabilidade e de imobilidade conduzem à vantagem competitiva sustentável?

Algumas pesquisas, com referência explícita ou não à Visão *Resource-Based*, buscaram tratar sobre a heterogeneidade dos recursos, e que as suas inimitabilidade e imobilidade, ainda que utilizando outros termos, conduziram a uma melhor performance, bem como resultar em vantagem competitiva sustentável. Anterior ao termo *Resource-Based*, Lippman e Rumelt (1982) ao proporem e testarem um modelo para tratamento analítico da ambigüidade causal, no qual conceituaram a imitabilidade incerta, evidenciaram, entre outros resultados, a diferença na eficiência das empresas nas indústrias que “apresentem” imitabilidade incerta. Por imitabilidade incerta, entende-se a obtenção da função de custo real de um recurso, apenas após a entrada no ambiente competitivo, tendo sido realizados investimentos não recuperáveis. Ao assumir esta incerteza na criação de novas funções de custos, ou de forma equivalente, novos recursos, explica-se a origem da diferença na eficiência das empresas. Outra correlação importante é a existente com a mobilidade dos recursos. Na ausência da imitabilidade incerta, um recurso único, ou recentemente criado, será facilmente “transportado” para o ambiente competitivo, pela reprodução de sua função de custo.

Cool e Dierickx (1989) propuseram uma abordagem na qual o acúmulo de ativos “não intercambiáveis”, como competências específicas, conhecimentos e valores, acumulados em decorrência do aprendizado pela execução e treinamento, e a sua aplicação aos produtos comercializáveis, conduz à vantagem competitiva sustentável. Uma vez que estes recursos não podem ser “comercializados”, o ambiente competitivo só terá acesso aos mesmos pela imitação, pelo acúmulo de recursos similares ou pela substituição. Analisam que o grau de imitabilidade destes recursos depende das seguintes características do processo de acumulação: o quanto à redução do tempo em que se completa o acúmulo conduz à redução dos ganhos, o quanto o processo de acumulação é dependente do “estoque anterior” ou de outros recursos, do grau de erosão dos recursos acumulados e da ambigüidade causal. Em resumo, os recursos acumulados se tornam “estratégicos” à medida que não são comercializáveis, inimitáveis e insubstituíveis.

Reed e DeFillippi (1990) associaram a obtenção de vantagem competitiva à criação de barreiras à imitação, cujas “alturas” decorrem do grau de ambigüidade causal, e que impedem o completo entendimento pelo ambiente competitivo de como esta vantagem é conquistada. Assinalam ainda que a vantagem competitiva sustentável é equivalente à manutenção da “altura” da barreira à imitação, frente à erosão pelo nível de competição, e que decorre do re-investimento nos fatores formadores da ambigüidade causal: o conhecimento tácito, a complexidade do inter-relacionamento entre recursos, competências e especificidade de ativos da organização.

Peteraf (1993) apresenta um modelo no qual quatro condições devem ser atendidas para atingir a vantagem competitiva sustentável. A primeira é a heterogeneidade dos recursos entre as empresas, permitindo que algumas apresentem recursos superiores a outras, obtendo conseqüentemente ganhos. Uma vez que estes recursos são limitados, esta vantagem competitiva só pode ser mantida se não ocorrer uma expansão descontrolada, isto é, promovida pela própria organização, ou a imitação pelo ambiente competitivo. A segunda, denominada limite à competição *ex post*, indica que uma vez obtida a vantagem competitiva, a mesma só será sustentável se existirem fatores que limitem a competição por esta

vantagem. Imitabilidade imperfeita, substituição imperfeita, mecanismos de isolamento (caput Rumelt, 1984), barreiras de mobilidade e ativos não comercializáveis são alguns dos fatores citados como limitadores da competição. A terceira é a mobilidade imperfeita dos recursos que conferem a vantagem competitiva. Uma vez que estes recursos permanecerão disponíveis e seus ganhos serão compartilhados pela empresa, podem então ser considerados fontes de vantagem competitiva sustentável. A quarta, denominada limite à competição *ex ante*, indica a necessidade da competição limitada para que possam ser desenvolvidos recursos com mobilidade imperfeita. Estas quatro condições apresentam forte correlação, sendo a heterogeneidade a condição necessária mais relevante.

Foss e Knudsen (2001) concordam que a heterogeneidade e a imobilidade dos recursos são fatores importantes, na análise da vantagem competitiva sustentável. Argumentam, entretanto, que as pesquisas e abordagens propostas, entre as quais destacam-se (BARNEY, 1991) e (PETERAF, 1993), não estabeleceram de forma clara se estas são condições necessárias e suficientes para a obtenção da vantagem competitiva sustentável. Sustenta que a heterogeneidade, “redefinida” como propriedades de eficiência diferencial dos recursos, é resultante da incerteza e da imobilidade. Define a incerteza como a “heterogeneidade” na expectativa e a vantagem competitiva sustentável como os lucros diferenciais, estritamente positivos, em excesso aos custos de oportunidade, incluindo o custo de capital, que são sustentados no equilíbrio. Para Foss e Knudsen a incerteza e a imobilidade são as duas únicas condições para a obtenção da vantagem competitiva sustentável.

Hall (1993) propõe uma abordagem para a identificação de fontes intangíveis de vantagem competitiva sustentável. Estas fontes intangíveis são as capacidades desenvolvidas pela empresa a partir de recursos intangíveis como patentes, a reputação de produtos, contratos e licenças; o conhecimento e a competência dos empregados, a cultura da organização etc. As capacidades desenvolvidas a partir destes recursos, certamente apresentarão imobilidade e inimitabilidade. Classifica as capacidades como regulatórias, de posicionamento,

estas baseadas nos ativos da empresa, funcionais e culturais, estas baseadas nas competências da empresa, e submete sua abordagem a um grupo de executivos de diferentes indústrias, obtendo como resultado, uma boa receptividade do método proposto na obtenção da vantagem competitiva sustentável.

A revisão dos trabalhos citados evidenciou que a exploração adequada de recursos inimitáveis e imóveis conduz à vantagem competitiva sustentável. Este resultado será utilizado para identificar o tipo de estratégia de acordo com a exploração de seus recursos que apresentem inimitabilidade e imobilidade.

2.1.2.3. A Visão *Resource-Based* e o Gerenciamento da Estratégia Corporativa

Estabelecidas as condições em que os recursos permitem a obtenção da vantagem competitiva sustentável, cabe analisar se uma abordagem baseada na Visão *Resource-Based* é adequada ao gerenciamento estratégico das empresas.

A Visão *Resource-Based* ao discutir sobre importância dos recursos, inimitáveis e imóveis, das empresas, incorpora conceitos centrais da pesquisa estratégica. Alguns destes conceitos são: o tipo e as fontes dos rendimentos das empresas, as competências distintivas, decorrentes da “posse” dos recursos em determinado instante, e da diversidade estratégica resultante, considerando-se o limite, a motivação e a direção do crescimento, e a influência sobre a performance são alguns destes conceitos (MAHONEY; PANDIAN, 1992). Prahalad e Hamel (1990) defendem o desenvolvimento de uma arquitetura estratégica para toda a corporação, estabelecendo objetivos voltados à construção da competência distintiva, resultante da alocação de recursos de forma consistente e do desenvolvimento de infra-estrutura administrativa adequada. Outros resultados são a criação de uma cultura gerencial, de espírito de equipe, capacidade de desafio, disposição ao compartilhamento de recursos, proteção das experiências e pensamento voltado para o longo prazo. Observa-se que os próprios resultados podem se converter em fonte de imitabilidade incerta e imobilidade, do ponto de vista do ambiente competitivo. Com foco na criação da vantagem competitiva pela

adequação da estratégia corporativa, Collis e Montgomery (1998) defendem que a estratégia corporativa de sucesso é aquela em que a exploração dos recursos está alinhada com os negócios da empresa e esta, organizada de forma a prospectar estes e outros recursos que levem à sustentação desta vantagem.

A revisão de abordagens que permitam a análise e contribuam para o gerenciamento estratégico das empresas apresenta-se relevante para justificar a aplicação da Visão *Resource-Based*. Neste sentido, duas propostas se destacam pela sua contribuição prática: a abordagem proposta por Grant (GRANT, 1991) e a abordagem VRIO (BARNEY, 2002, p.158-185). A primeira foi proposta com o objetivo de contribuir para a integração das estratégias corporativa e de negócios. Na estratégia corporativa, o papel dos recursos é determinar as fronteiras geográficas e industriais das atividades das empresas, enquanto que a estratégia de negócios é explorada a relação entre recursos, competição e lucratividade. A formação da estratégia é realizada em cinco processos: (1) a análise dos recursos da empresa; (2) a avaliação das capacidades da empresa; (3) a análise do potencial de lucro dos recursos e capacidades; (4) elaborar a estratégia adequada aos recursos e capacidades da empresa e (5) reavaliar a estratégia, realizando um processo de realimentação, com o objetivo de estender e aprimorar os recursos e capacidades e obter uma vantagem competitiva sustentável. A figura 1 ilustra esta abordagem (GRANT, 1991).

Um dos aspectos interessantes desta abordagem está na visão dos recursos e capacidades como orientadores da direção estratégica a ser adotada pela empresa. Em resumo, a definição dos negócios em termos do que se é capaz de executar pode oferecer uma base mais duradoura para o gerenciamento estratégico da empresa, do que uma definição baseada nas necessidades que o negócio procura satisfazer. Além disso, reforça a importância da incerteza e da imobilidade dos recursos para a vantagem competitiva sustentável, em perspectivas que permitem grande contribuição na identificação de *construtos* para análise da estratégia das empresas.

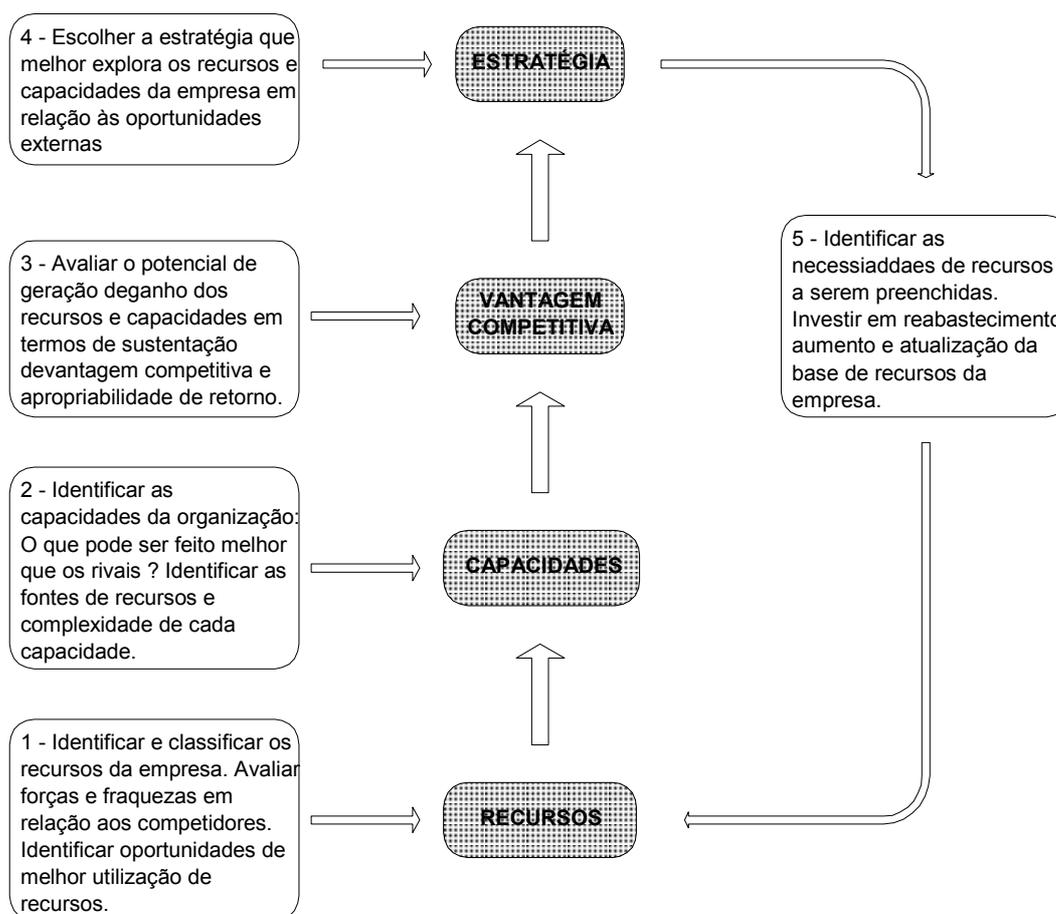


Figura 1 – Abordagem de avaliação estratégica proposta por Grant (GRANT, 1991).

A abordagem VRIO – *Valuable, Rare, Inimitable e Organization* (BARNEY, 2002, p.158-185) – apresenta-se como uma ferramenta para tornar mais tangível a premissa de heterogeneidade, esta discutida e revista como resultado da incerteza e imobilidade dos recursos. Para obter a vantagem competitiva, os recursos da empresa devem apresentar valor e raridade no ambiente competitivo, devem ser difíceis de copiar, e a empresa que os possui deve estar estruturada para aproveitar estas características. Para a condução da análise da empresa com base na visão *Resource-Based* são propostas as seguintes questões: (1) a questão do valor: os recursos e capacidades da empresa permitem-na responder às ameaças e oportunidades do ambiente competitivo? (2) a questão da raridade: os recursos de valor são controlados por poucos concorrentes? (3) a questão da inimitabilidade: as empresas que não possuem estes recursos

enfrentarão desvantagens de custo para copiá-los ou desenvolvê-los? (4) a questão da organização: as políticas e processos da empresa são adequados para garantir a exploração dos recursos de valor, raros e difíceis de copiar? . O quadro 1 ilustra a correlação das perguntas acima com a performance econômica, a vantagem competitiva e as forças e fraquezas das empresas.

Valor?	Raro?	Difícil de Imitar?	Explorado pela empresa?	Implicações Competitivas	Performance Econômica	Força ou Fraqueza
Não	-	-	Não	Desvantagem competitiva	Abaixo do Normal	Fraqueza
Sim	Não	-	Sim	Paridade competitiva	Normal	Força
Sim	Sim	Não	Sim	Vantagem Competitiva Temporária	Acima do Normal	Força e competência distintiva
Sim	Sim	Sim	Sim	Vantagem competitiva sustentada	Acima do Normal	Força e competência distintiva sustentável

Quadro 1 – Correlação dos parâmetros VRIO com a performance econômica, a vantagem competitiva e as forças e fraquezas das empresas.

Os aspectos a se destacar na abordagem VRIO dizem respeito a implicações mais abrangentes da visão *Resource-based*. A extensão da responsabilidade pelo sucesso na obtenção da vantagem competitiva para toda a empresa, e não somente a nível gerencial; a dificuldade na implantação da estratégia; a questão dos recursos socialmente complexos; a adequação da estrutura, valores, políticas e processos, para garantir a busca pela vantagem competitiva sustentável.

Os conceitos discutidos acima indicam aspectos complementares e suplementares da visão *Resource-Based* aos modelos estratégicos, de análise e gerenciamento, baseados no ambiente. Este aspecto também pode ser considerado

no que diz respeito à pesquisa das organizações industriais, abrindo uma porta para a integração da visão *Resource-Based* com a questão da identificação de grupos estratégicos (MAHONEY ; PANDIAN, 1992).

2.1.3. Grupos Estratégicos e a Equivalência Estratégica

Esta seção revisa o conceito grupos estratégicos através da seguinte abordagem: (1) discussão do conceito de grupos estratégicos, (2) as questões da performance e rivalidade internas e entre os grupos estratégicos e de sua dinâmica e (3) a delimitação dos grupos estratégicos através das barreiras de entrada, de saída e de mobilidade.

2.1.3.1. O Conceito Grupo Estratégico

O conceito de grupos estratégicos foi proposto por Hunt em 1972 para apresentar uma explicação para a diferença de performance, evidenciada em pesquisa realizada na indústria de eletrodomésticos da linha branca dos Estados Unidos na década de sessenta (McGEE; THOMAS, 1986). Foram identificados quatro “grupos” que pareceram estar seguindo estratégias diferentes. A definição proposta por Hunt foi:

“Um grupo de empresas de uma determinada indústria altamente equivalentes... com relação à estrutura de custo, grau de diversificação de produtos... organização formal, sistemas de controle, gerenciamento de recompensas e punições... e visões pessoais e preferências para possíveis resultados...” (THOMAS; VENKATRAMAN, 1988).

Além deste conceito, tentou-se isolar e descrever as “barreiras de entrada” características de cada grupo. Uma outra definição para grupos estratégicos é apresentada por Barney (BARNEY, 2002). Grupos estratégicos são subconjuntos de empresas de uma indústria que “enfrentam” ameaças e oportunidades semelhantes, diferentes de outras ameaças e oportunidades, “enfrentadas” por outras empresas na mesma indústria. Embora pareçam diferentes estas definições

permitem a realização da análise do ambiente competitivo de um setor, com base nas estratégias evidenciadas, seja para enfrentar as ameaças e aproveitar as oportunidades, seja para ajustar a empresa para obter e manter determinada performance. Assim, para efeito de discussão do gerenciamento das estratégias das empresas, à luz do conceito de grupos estratégicos, é adotada a seguinte definição: grupos de empresas de uma mesma indústria que apresentem equivalência no gerenciamento estratégico em uma arena competitiva.

Desde sua formulação, este conceito é objeto de pesquisas teóricas e empíricas que buscam confirmar sua existência, sua contribuição à avaliação da performance e à formação das estratégias das empresas. Estas pesquisas sobre grupos estratégicos apresentaram limitações como a aceitação implícita de limites “bem-definidos” e pré-estabelecidos das indústrias – limitadas às fronteiras nacionais – falta de clareza nas descrições dos grupos identificados, ausência de uniformidade na escolha das variáveis utilizadas para evidenciar a semelhança estratégica e fraca evidência de variações de performances intra e inter grupos (THOMAS; VENKATRAMAN, 1988). A esta lista, pode ser acrescido o fato de poucos estudos, evidenciarem a relevância, verificada pelo conhecimento e entendimento do setor, das variáveis utilizadas para a indústria pesquisada (McGEE; THOMAS, 1986).

Buscando propor uma abordagem alternativa, Reger e Huff realizaram pesquisa – na indústria bancária da região de Chicago, Estados Unidos no período de 1982 a 1985 – utilizando a percepção cognitiva dos estrategistas sobre eventuais grupamentos de estratégias semelhantes para identificar os grupos estratégicos. Neste trabalho, duas proposições são apresentadas: (1) a percepção dos estrategistas sobre a estratégia dos competidores tenderá a ser pelo agrupamento de características semelhantes, e não pela percepção de uma estratégia única para cada competidor, ou percepção de uma estratégia única da indústria; (2) esta estrutura de grupos é compartilhada pelos estrategistas da indústria, e não uma percepção diferente por parte de cada estrategista.

Na metodologia empregada, os estrategistas participantes foram convidados a organizar o ambiente da indústria de acordo com as estratégias percebidas, e utilizando o conjunto de variáveis que achasse adequado. As proposições apresentadas foram confirmadas, embora algumas empresas não apresentaram estratégias consistentes no período, enquanto outras apresentaram estratégias que não puderam ser explicitadas com base nas principais dimensões estratégicas propostas pelos estrategistas. Com base nestes resultados, entre outros, propõem que as empresas que “constituíntes” de um grupo sejam classificadas em: (1) empresas núcleos, aquelas cujas estratégias caracterizam o grupo; (2) empresas secundárias, cujas estratégias são menos consistentes que as das empresas núcleo, porém com algum grau de semelhança; (3) empresas transitórias, nas quais são evidenciadas mudanças de estratégia de um grupo para outro, e ainda que esta “adesão” a um tipo, ou ao outro, apresente graus de “aderência”. Outro resultado importante, está no fato desta abordagem permitir a existência de interseções entre os diversos grupos identificados, o que permitiria uma visão mais realista do ambiente competitivo de um setor. Como as outras abordagens, esta também apresenta limitações como não evidenciar momentos de mudança ou apresentar viés nos grupamentos encontrados, como decorrência das estruturas cognitivas dos estrategistas (REGER; HUFF, 1993).

Apesar das limitações apresentadas, os resultados das pesquisas realizadas com grupos estratégicos permitiram evidenciar a inadequação de uma visão homogênea de uma indústria, na avaliação da performance, a partir de indicadores financeiros e operacionais, e sobre a evolução das empresas em suas arenas competitivas (THOMAS; VENKATRAMAN, 1988). O reconhecimento desta heterogeneidade implica no “agrupamento” de empresas que realizaram opções estratégicas semelhantes. Desta forma, o conceito de grupo estratégico, complementado pelas pesquisas com abordagens “multidisciplinares” envolvendo organizações industriais, estratégia de marketing, comportamento administrativo e gerenciamento estratégico, enfatiza a empresa como importante unidade de análise e retorna para a arena competitiva o centro das decisões estratégicas (McGEE; THOMAS, 1986).

A utilização da empresa como unidade de observação pela análise dos grupos estratégicos, apresenta interessantes paralelos com a teoria de crescimento da empresa como inicialmente proposta por Downie, Penrose e Marris no fim da década de 50 (McGEE; THOMAS, 1986). Suas contribuições foram, entre outras: a de passar a enxergar a empresa como um amplo conjunto de recursos (WERNERFELT, 1984), as implicações competitivas de recursos produtivos inelásticos como equipes gerenciais, alta direção e capacidades empreendedoras (BARNEY, 2002, p.155), a vinculação do crescimento da empresa com a lucratividade e sua importância no processo competitivo, a de que determinadas características organizacionais e gerenciais, mais tarde convenientemente denominadas competências distintivas, facilitam o sucesso das iniciativas estratégicas e o subsequente desenvolvimento de estruturas corporativas (McGEE; THOMAS, 1986). Tais conceitos estão altamente correlacionados com os conceitos da Visão *Resource-Based*.

2.1.3.2. Grupos Estratégicos: Dinâmica, Performance e Rivalidade

Três aspectos merecem destaque no estudo de grupos estratégicos. A sua dinâmica, isto é, a evolução dos grupos ao longo do tempo, a performance e a rivalidade das empresas entre membros de um mesmo grupo, e de grupos diferentes.

A dinâmica dos grupos estratégicos reflete o próprio desenvolvimento das estratégias das empresas na arena competitiva, no sentido de corrigir o posicionamento estratégico. Poucas pesquisas entretanto buscaram avaliar este aspecto, talvez em função da dificuldade em resgatar as informações de períodos passados, minimizada pelo uso de bancos de dados. Quando é possível dispor de um banco de dados, a dificuldade que se coloca é a adequação das informações disponíveis às variáveis utilizadas para identificação dos grupos. De uma forma geral esta dinâmica pode ser vista pelos seguintes aspectos: (1) os efeitos no grupo, mudança da estratégia do grupo todo, rearranjo dos grupos em decorrência de similaridade estratégica e alteração no número de grupos, pela mudança de estratégia de alguns de seus membros; (2) como se desenrola esta mudança no

ambiente dos grupos e das empresas. Para avaliar o primeiro aspecto, uma pesquisa procurou identificar o comportamento dos grupos estratégicos em períodos de estabilidade, crescimento e declínio na indústria mundial de perfuração de petróleo. Foram identificados quatro grupos estratégicos na indústria utilizando-se variáveis de relevância para o setor, estabelecidas a partir de entrevistas com executivos de empresas perfuradoras, e avaliados os dados dos anos 1966 – com 41 empresas – 1973 – com 49 empresas – 1981– com 110 empresas – e 1984, com 123 empresas (MASCARENHAS, 1989). As principais conclusões foram:

- as mudanças na estratégia característica do grupo estavam associadas a mudanças ambientais de crescimento e declínio, caracterizando uma perspectiva de adaptação a estas mudanças, porém sendo preservada a escolha estratégica que caracterizava o grupo;
- as mudanças de membros entre os grupos foram mais acentuadas em períodos de declínio, tendo ocorrido mais entre grupos similares. No período de estabilidade podem ser associadas poucas atividades de pesquisa e pequena mobilidade, enquanto que nos períodos de crescimento, ainda que identificadas mais entradas por parte de empresas de fora da indústria, pequena mobilidade entre grupos foi evidenciada;
- após longo período de estabilidade, foi identificado o aparecimento de um novo grupo. Este novo grupo foi “criado” por um líder de um dos grupos, em resposta ao aumento da competição interna e externa aos grupos, não tendo sido a nova escolha estratégica acompanhada pelos outros membros, seja por relutância em assumir novo posicionamento ou por impossibilidade em imitar a nova estratégia;
- o ambiente e a estrutura da indústria não foram os únicos fatores determinantes das mudanças estratégicas. As características das empresas e a escolha estratégica apresentaram papel importante.

O segundo aspecto foi avaliado por uma pesquisa sobre a estabilidade do arranjo de grupos estratégicos, associada ao número de grupos e seu posicionamento ao longo do tempo, a correlação entre o processo de adaptação da estratégia e o movimento dos grupos, e da mobilidade entre grupos estratégicos na indústria de seguros dos Estados Unidos no período 1970 a 1984, divididos em

nove intervalos. Foram utilizados os dados das 30 maiores empresas do setor, correspondendo a cerca de 75% do mercado no período, em cada um dos intervalos. sendo identificados inicialmente 5 grupos, chegando ao final do período de avaliação com 6 grupos. Foram observados o aparecimento, máximo registrado, de 3 grupos estratégicos no período 1974-75, e o desaparecimento 2 grupos nos períodos 1973, 1976 e 1983-84, ao longo do período de estudo (FIEGENBAUM; THOMAS, 1993). As principais conclusões foram:

- os grupos estratégicos são elementos estáveis da estrutura da indústria, tendo sido observados três grupos estáveis, cujas estratégias adotadas estavam consistentes com as variações observadas nas dimensões estratégicas escolhidas para evidenciar os grupos;
- as mudanças das estratégias das empresas ocorreram de forma “quântica”, isto é, só foram realizadas quando a inadequação da estratégia em relação ao ambiente atingiu um limite crítico, talvez insustentável, na avaliação das empresas. Até este ponto, estas empresas experimentaram poucas mudanças em suas estratégias.
- as mudanças de grupos estratégicos foram raras, sendo as barreiras econômicas e organizacionais responsáveis pela manutenção das posições do grupo mais forte. A existência de baixos níveis de mudança é explicada pelo fato dos grupos estratégicos apresentarem-se com fortes pontos de referência do comportamento das empresas, atuando no sentido de reforçar as “normas” dos grupos e das barreiras de mobilidade.

Os resultados apresentados por estas pesquisas evidenciam uma estabilidade nos grupamentos de empresas, permitindo que, para efeito de caracterizar as estratégias de um período pequeno, a identificação de grupos estratégicos em uma indústria seja uma abordagem útil para o gerenciamento estratégico. Outro fator importante são as barreiras de mobilidade, relevantes na preservação da estrutura de grupamentos de uma indústria.

Embora a performance, entre membros de um determinado grupo estratégico ou de grupos diferentes, seja um tema amplamente discutido nas pesquisas sobre grupos estratégicos, não existe uma conclusão definitiva sobre esta questão. A controvérsia varia desde os indicadores para avaliar esta performance –

normalmente lucratividade, em alguma de suas formas como retornos sobre ativos, e participação no mercado – até as variáveis utilizadas. Na avaliação de pesquisas empíricas sobre grupos estratégicos, as pesquisas não conseguiram evidenciar correlação conclusiva entre pertencer a um determinado grupo estratégico e apresentar então, uma melhor performance (THOMAS; VENKATRAMAN, 1988). Entretanto, algumas pesquisas buscaram avaliar a questão da performance a partir de uma abordagem indireta.

Uma delas associa a lucratividade das empresas da indústria farmacêutica dos Estados Unidos, nos períodos 1963-69, 1970-74, 1975-79 e 1980-82, à rivalidade entre os membros de um mesmo grupo e entre grupos. Nesta pesquisas foram inicialmente identificados 6 grupos, que ao longo do período avaliado apresentaram variações de 4,5 e 6 grupos nos períodos, respectivamente, 1975-79, 1970-74 e 1980-82. Foi evidenciado o aumento da rivalidade entre os membros de um mesmo grupo, bem como o aumento da rivalidade entre os grupos, pela observação de um re-posicionamento estratégico, tendo sido estas rivalidades consideradas responsáveis pela diminuição da lucratividade da indústria. Outra consequência diz respeito às mudanças de empresas de seu grupo original, período 1963-69, para outros grupos, observadas as posições no período 1980-82. Foi também evidenciado, em alguns períodos, que as empresas que apresentaram maior lucratividade, também detinham a maior participação de mercado (COOL; DIERICKX, 1993).

Outra pesquisa, utilizando a mesma base de dados da pesquisa citada acima, avaliou a performance através da diferença de riscos dos investimentos estratégicos, obtido a partir do resultado sobre as vendas. O principal argumento para utilização deste indicador está no fato de que os investimentos estratégicos apresentam uma relação risco-retorno correlacionada aos fatores: ativos, os recursos e competências, acumulados pelas empresas e das imperfeições do mercado. Esta correlação reside no fato, de que as empresas irão responder de formas diferentes às perturbações, decorrentes das imperfeições do mercado, em suas estratégias. Esta diferença deve-se à heterogeneidade nos ativos acumulados. As conclusões do estudo foram:

- as empresas pertencentes ao mesmo grupo estratégico apresentam riscos diferentes;
- as empresas pertencentes ao mesmo grupo estratégico apresentam retornos diferentes;
- as relações risco-retorno não apresentaram estabilidade ao longo do período avaliado, tendo sido positivamente correlacionadas nos períodos 1963-69 e 1970-74, ligeiramente negativa no período 1975-79 e significativamente negativa no último período.

Outro aspecto interessante deste estudo foi avaliação da correlação risco-retorno como resultante da “convivência” no mesmo grupo de empresas que obtém sucesso e que apresentam problemas na implementação desta estratégia, caracterizada como correlação risco-retorno negativa, e dos resultados dos investimentos estratégicos destas empresas, caracterizadas como correlação risco-retorno negativa (COOL; SCHENDEL, 1988).

2.1.3.3. A Delimitação dos Grupos Estratégicos

Assumir que os membros de um mesmo grupo estratégico apresentam equivalência no gerenciamento estratégico implica nas questões: (1) de que forma se pode identificar esta equivalência? (2) esta equivalência estratégica compreende atingir os mesmos objetivos? A busca pelas respostas a estas questões está implícita, nos trabalhos de pesquisa sobre grupos estratégicos, através das variáveis utilizadas, nos objetivos, na metodologia empregada e em suas conclusões. Com base nas considerações sobre estas pesquisas feitas por Thomas e Venkatraman (1988), McGee e Thomas (1986) e nos aspectos da dinâmica e performance em grupos estratégicos, o conceito de barreiras de mobilidade é discutido com o objetivo de caracterizá-lo como elemento de delimitação dos grupos estratégicos, e conseqüentemente, de identificação da equivalência estratégica.

Os conceitos de barreiras de entrada e de saída de uma indústria “atribuem” a mesma uma homogeneidade que não corresponde à realidade. Economia de escala, políticas de governo e compromissos com fornecedores e clientes, apenas para citar alguns exemplos de barreiras de entrada e de saída, são vistos como

“ativos” coletivos, protetores dos lucros das empresas, dos competidores externos a indústria. Entretanto, a heterogeneidade existente nas indústrias leva a conclusão que os investimentos na manutenção destas barreiras será distinto para cada empresa, e decorrente da escolha do posicionamento estratégico de cada uma, resultando em vulnerabilidades ao longo desta barreira. A identificação destas vulnerabilidades por parte de novos entrantes irá orientar sua mobilidade no sentido de ter acesso à indústria onde forem menores as “alturas” das barreiras de entrada e saída. Tal raciocínio permite a ampliação do conceito de barreiras de entradas e de saídas para o conceito de barreira de mobilidade de novos entrantes do ambiente competitivo. A observação da heterogeneidade da indústria conduz também a identificação da criação de barreiras de mobilidades internas à indústria como decorrência das escolhas estratégicas das empresas. Estas barreiras de mobilidade resultantes da forma como cada empresa, ou grupos de empresas de uma indústria, direcionam seus recursos com o objetivo de estabelecer barreiras de entradas ou de saídas. Os grupos estratégicos surgiriam então, pela similaridade nos investimentos e ações na construção de barreiras de proteção ao seu mercado (CAVES; PORTER, 1977).

O conceito de “barreira de mobilidade”, representando os obstáculos existentes à movimentação de concorrentes, internos e externos, de uma indústria, apresenta-se então como um orientador para a identificação dos grupos estratégicos. Com o objetivo de demonstrar o valor desta abordagem, uma pesquisa foi realizada na indústria mundial de perfuração de petróleo em 1982, tendo também como objetivos secundários avaliar a correlação das barreiras de mobilidade com a lucratividade (MASCARENHAS; AAKER, 1989). As variáveis utilizadas na identificação dos grupos estratégicos foram baseadas nas características consideradas barreiras de mobilidade na indústria, na opinião de executivos, especialistas financeiros e consultores do setor. Foram identificados três grupos estratégicos na indústria, de um conjunto de 679 empresas. As principais conclusões foram:

- a avaliação de barreiras de mobilidade permite a identificação de ativos e competências que distinguem as empresas de uma indústria, re-orientando a

descrição das empresas da postura “o que se faz” para a postura “o que se é (ativos e competências)”;

- a estrutura dos grupos estratégicos mostrou-se estável, tendo sido identificados apenas dois casos, em 109 potenciais, de superação da barreira de mobilidade existente, evidenciando a força destas barreiras nesta indústria;
- os grupos estavam melhores protegidos por barreiras de mobilidade apresentaram menor lucratividade, explicada pela forte correlação de aspectos relevantes, como poder dos clientes e agressividade dos investimentos, quando as taxas de juros e os custos dos equipamentos estavam altos; além disto, sugeriu que as barreiras de saída podem se constituir em armadilhas em condições de dificuldades econômicas, levando à redução dos lucros;
- os grupos aparentemente menos protegidos apresentaram barreiras de mobilidade difíceis de imitar, como estilo gerencial e equipamentos depreciados;
- a evidência de assimetrias nas barreiras de mobilidade da indústria, desafiando os estrategistas a criar barreiras de “alturas” adequadas ao potencial dos competidores e redução das “barreiras de saída”; O entendimento desta assimetria como fator importante na compreensão e influência no ambiente competitivo da indústria.

Um aspecto curioso citado na pesquisa é a observação de um “fator de orgulho”. Os executivos dos dois grupos constituídos pelas empresas de maior porte não demonstraram interesse em mover-se na direção do grupo das empresas de menor porte, pelo fato de terem emergido deste terceiro grupo. Tal fator certamente influencia na elaboração da estratégia, constituindo-se em uma barreira de mobilidade, estabelecida de forma tácita, pelos próprios concorrentes deste terceiro grupo (MASCARENHAS; AAKER, 1989).

Outra pesquisa, ainda que seu objetivo principal fosse o de evidenciar uma performance superior com grupos estratégicos e fatores que explicassem esta performance, na indústria de perfumes, cosméticos e assemelhados dos Estados Unidos no período 1973-84, evidenciou a importância das barreiras de mobilidade,

e sua assimetria, como fatores de manutenção da vantagem competitiva sustentada (OLUSOGA; MOKWA; NOBLE, 1995). As principais conclusões relativas a este ponto foram:

- as barreiras de mobilidade relevantes na indústria são a diferenciação de produtos e os recursos de capital, tendo sido evidenciado que as empresas que apresentaram melhor performance, investiram nos fatores construtores destas barreiras;
- embora o tamanho das empresas não tenha se mostrado como único determinante da performance, grandes investimentos na construção das barreiras de mobilidade estavam relacionados às maiores empresas da indústria;
- o aumento de número de grupos estratégico evidenciou uma assimetria na barreira de mobilidade de diferenciação, característica marcante dos grupos emergentes;
- o re-direcionamento dos recursos de capital para fatores de longo prazo – como políticas de distribuição e investimentos em fábricas – para fatores de curto prazo, embora os gastos com propaganda mantiveram forte utilização de recursos.

As considerações destas pesquisas reforçam as conclusões do trabalho de McGee e Thomas (1986), no que diz respeito ao estabelecimento de barreiras de mobilidade pelo aproveitamento de recursos como: a estrutura organizacional, sistemas de controle, a diversificação e da integração, e as relações com outros grupos, como por exemplo, união com outras empresas, grupos de consumidores e reguladores. Estas barreiras são essencialmente limites à replicação e à imitação, atuando como inibidores do movimento das empresas de uma posição estratégica para outra.

A discussão conduzida sobre as barreiras de mobilidade, com base nas pesquisas apresentadas, permite caracterizá-las como delimitadoras dos grupos estratégicos. Decorrente da orientação de recursos das empresas na manutenção de seu posicionamento estratégico, a assimetria nestas barreiras permite delinear fatores e recursos que caracterizariam uma equivalência estratégica, bem como os

objetivos destas, na busca pela vantagem competitiva sustentada. A figura 2 apresenta uma visão pictórica da “partição” do ambiente competitivo pelos grupos estratégicos, sendo as divisas e fronteiras com outras indústrias estabelecidas pelas barreiras de mobilidade e de entrada / saída.

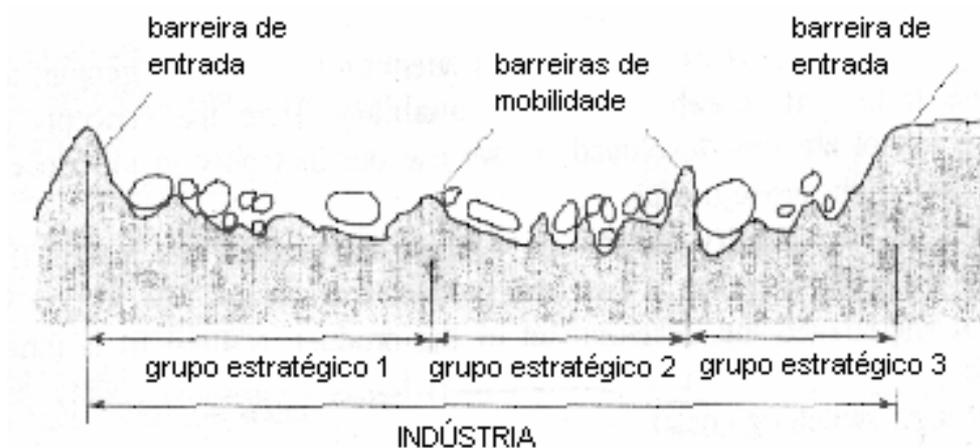


Figura 2 – Visão pictórica do ambiente competitivo de uma indústria.

2.1.4. A Visão *Resource-Based* e a Delimitação de Grupos Estratégicos

Discutidos os aspectos básicos sobre a visão *Resource-Based* e os grupos estratégicos cabe então a questão: de que forma os conceitos inimitabilidade e imobilidade dos recursos da visão *Resource-Based* são adequados à identificação de grupos estratégicos? A resposta a esta questão é o objetivo principal deste trabalho, cabendo, entretanto, algumas considerações sobre as contribuições que a aplicação sinérgica da visão *Resource-Based* e o conceito de grupos estratégicos, podem trazer para o gerenciamento estratégico.

Ao sugerir a integração da visão *Resource-Based* com a análise de grupos estratégicos, Mahoney e Pandian (1992) defendem a opinião de McGee e Thomas (1986) da existência de paralelos com a teoria de crescimento das empresas, articuladas por Downie, Penrose e Marris, e a análise de grupos estratégicos, ainda que utilizando diferentes grupos de análise, propondo a questão: Podem os

recursos raros e inimitáveis ser fonte de vantagens competitivas dos grupos estratégicos ?

Dess e Davis (1984) ao realizarem estudo no qual buscaram associar a performance das empresas ao fato de pertencer ou não a grupos estratégicos, pré-determinados, e utilizando três das estratégias genéricas de Porter, chamaram a atenção para as diferenças entre as estratégias intencionais e as realizadas. Estas podem resultar da inabilidade das empresas em traduzir em ações as suas intenções estratégicas devido a ocorrência, isolada ou em conjunto, dos fatores: mudanças inesperadas no ambiente, a inexistência de capacidade de implementação ou expectativas irreais. A contribuição da visão *Resource-Based* para minimizar esta diferença, está no fato de que a utilização desta abordagem na condução da estratégia pelas empresas não está relacionada apenas com a aplicação dos recursos existentes, mas também com o desenvolvimento e manutenção dos “estoques” de recursos, com o objetivo de sustentar e ampliar as posições de vantagem competitiva (GRANT, 1991). A capacidade em desenvolver e manter recursos de valor estratégico também pode contribuir para a avaliação de estratégias alternativas e de caracterização das diferenças entre os membros de um mesmo grupo.

Cool e Schendel (1988) argumentam que uma das condições para uma estratégia de negócios eficiente é a adequação dos ativos acumulados, recursos e competências, às ações destinadas a explorar uma oportunidade no mercado. Se a estratégia da empresa é incongruente com seus recursos acumulados, provavelmente será menos eficiente que outra empresa que adote estratégia semelhante, mas cuja adequação das iniciativas estratégicas aos ativos acumulados seja maior. A estratégia adotada na alocação dos recursos adquire importância, na medida que altera as reservas dos ativos acumulados. O impacto provocado, entretanto, na vantagem competitiva é normalmente indireto, isto é, resulta de variações incrementais nas reservas de competências e / ou ativos acumulados. Sendo assim, nos limites de um mesmo grupo estratégico, as empresas com inadequação entre sua estratégia e seus recursos, provavelmente apresentarão

maior exposição ao risco em seu investimento estratégico, do que os outros membros do grupo com maior adequação entre estratégia e recursos.

Em seu trabalho sobre a perspectiva cognitiva dos grupos estratégicos, Reger e Huff (1993), ao apresentar propostas para futuras pesquisas, defendem o potencial dos grupos estratégicos na avaliação de estratégias alternativas. Os desvios em relação à empresa núcleo por parte das empresas secundárias poderiam indicar a “incapacidade” destas em implementar as estratégias da primeira, ou uma tentativa deliberada no sentido de diferenciar-se das empresas núcleos. A distinção entre imitação e diferenciação estaria associada às diferenças existentes nas características da empresa núcleo e das empresas secundárias. Neste caso, a avaliação das características da empresa núcleo a luz dos conceitos de inimitabilidade e da imobilidade dos recursos poderia evidenciar as dificuldades de imitação. Uma avaliação de seus próprios recursos permitiria às empresas secundárias estabelecerem as estratégias de diferenciação que as levaria à obtenção de vantagem competitiva.

A principal questão desta sinergia entre a visão *Resource-Based* e os grupos estratégicos a ser explorada por este trabalho, é a da delimitação dos grupos estratégicos pelas barreiras de mobilidade. Como já discutido anteriormente, as barreiras de imitação atuam como inibidores do movimento das empresas de uma posição estratégica para outra (McGEE; THOMAS, 1986), constituindo-se em barreiras de mobilidade, estabelecendo assim as “fronteiras” dos grupos estratégicos. Tais barreiras à imitação seriam decorrentes da inimitabilidade e imobilidade resultantes da forma como é conduzida a exploração de recursos estratégicos das empresas. Identificar portanto, os graus de inimitabilidade e de imobilidade dos recursos pode ser uma abordagem a ser aplicada na identificação de grupos estratégicos.

2.2. MÉTODOS ESTATÍSTICOS E TÉCNICAS DE INTELIGÊNCIA COMPUTACIONAL APLICADAS A CLUSTERIZAÇÃO

2.2.1. Métodos Estatísticos – Análise de Clusters

As pesquisas sobre grupos estratégicos, sua identificação e análise, utilizam vários métodos estatísticos no tratamento dos dados obtidos de empresas. Uma amostra de alguns destes métodos empregados é apresentada em (THOMAS; VENKANTRAMAN, 1988), correlacionando estes métodos com a classificação, sugerida no mesmo trabalho, para as pesquisas de grupos estratégicos. Nesta classificação, quatro tipos de pesquisas são identificados: (I) avaliação da estratégia está concentrada em uma área funcional ou dimensão e as características dos grupos são estabelecidas à priori, sendo a análise dos dados utilizada para validar a classificação estabelecida teoricamente; (II) concentração em uma área funcional ou dimensão, porém obtendo-se os grupamentos de forma empírica, sobre uma base de dados específica; (III) avaliação da estratégia pela análise de múltiplas áreas funcionais, ou dimensões, e grupamentos estabelecidos previamente; (IV) múltiplas áreas funcionais ou dimensões e grupamentos obtidos empiricamente. A condução da discussão até o momento, sobre ambiente competitivo, a visão *Resource-Based* e grupos estratégicos, classifica esta dissertação como tipo III ou IV. A observação dos tipos de análise de dados empregados nas pesquisas, igualmente classificadas em (THOMAS; VENKANTRAMAN, p. 542-545, 1988), indicam a análise de clusters como a técnica mais utilizada. Com base nestes dados, são apresentados alguns aspectos desta técnica com o objetivo de permitir a sua comparação com a técnica proposta neste trabalho, utilizando-se inteligência computacional.

De uma forma geral, o termo Análise de Clusters serve para designar uma variedade de procedimentos utilizados para realizar uma classificação de dados. Especificamente falando, um método de clusterização é um procedimento estatístico multivariável que, busca reorganizar um conjunto de dados representativos de uma amostra, em grupos, os clusters, relativamente homogêneos. Estes grupos, e sua análise, têm quatro grandes aplicações: desenvolvimento de uma tipologia ou classificação, investigação abordagens conceituais de grupamento de entidades, geração de hipóteses através da

exploração de dados e o teste de hipóteses, pela verificação de classificação obtida por outros processos em um conjunto de dados. A primeira aplicação é provavelmente a de maior ocorrência, embora em muitos casos de análise de dados aplicada, a combinação destas aplicações é empregada para formar a base do estudo (ALDENDERFER; BLASHFIELD, 1984).

Um conceito importante da análise de clusters é o de similaridade, bem como os procedimentos para a sua medida. Apesar da aparente simplicidade, este conceito, e principalmente os procedimentos utilizados para medi-la, não pode se basear apenas no simples reconhecimento de que dois “objetos” são semelhantes ou não. Para que este conceito possa ser aplicado em pesquisas científicas, deve ser baseado em objetividade e procedimentos reproduzíveis. Nas pesquisas de características quantitativas, estas premissas são bem resolvidas pelo conceito de métrica, matematicamente representada pela Simetria, Desigualdade Triangular, Distinção de não Idênticos e Não Distinção de Idênticos, e pela aplicação de procedimentos estatísticos. Nos estudos qualitativos o conceito de similaridade pode ser baseado na comparação de características ou no grau de relacionamento existente entre as entidades comparadas. Desta forma, para aproveitar a potencialidade da análise de cluster deve-se estar atento para a escolha da medida de similaridade. Esta escolha deve estar embutida no projeto de pesquisa, sendo determinada pelo contexto prático, teórico e filosófico do problema de classificação.

Outra característica importante a ser levado em conta é a escolha das variáveis. O principal aspecto a ser observado é a escolha dentro de um contexto da teoria utilizada como base da classificação, para evitar uma “ingenuidade empírica“, isto é, a coleção e análise dos dados na esperança de que uma estrutura irá emergir apenas pela coleta de dados suficientes. Apenas como citação, as principais medidas de similaridade são: Coeficientes de Correlação, Medida de Distância, na qual se destacam a Euclidiana, a de Manhattan e a de Mahalanobis D^2 , a Associação de Coeficientes e a Similaridade Probabilística de Coeficientes. Estas medidas, bem como o conceito de similaridade, são aplicados na obtenção de

métodos que permitem a implementação de métodos para a análise de clusters (ALDENDERFER; BLASHFIELD, 1984).

Os métodos de clusterização podem ser classificados em sete famílias: Hierárquico por Aglomeração, Hierárquico por Divisão, Partição Iterativa, Pesquisa por Densidade, Analítica de Fatores, Acumulativa e Gráfica Teórica. Cada uma destas famílias representa uma perspectiva diferente na criação de grupos, podendo fornecer resultados diferentes para um mesmo conjunto de dados. O principal aspecto a ser ressaltado é que o método a ser adotado seja compatível com a natureza da classificação desejada, com as variáveis utilizadas e a medida de similaridade adotada para estimar a semelhança entre os objetos a classificar. Abaixo são apresentados os principais características das famílias mais utilizadas, com destaque para suas características relevantes.

- Hierárquico por Aglomeração: de simples entendimento e representação gráfica, uso de dendogramas, busca seqüencialmente criar grupos de acordo com seus valores de similaridade, de acordo com regras pré-estabelecidas. Os diversos métodos desta família distinguem-se basicamente pelas regras utilizadas. Apenas um processamento dos dados e diferentes soluções em função da ordem inicial são as principais desvantagens destes métodos. Os métodos mais conhecidos são: Conexão simples, Conexão completa, Conexão por média e Método de Ward.
- Partição Iterativa: através de uma partição inicial dos dados em clusters, realiza iterações consecutivas, até que não ocorram mais mudanças de grupos entre os dados. Três aspectos distinguem os métodos desta família: a partição inicial, isto é, a estimativa ou escolha direcionada dos centróides dos clusters, a forma como os dados são associados aos clusters, sendo *K-means* e *Hill climbing* os tipos básicos existentes, e os critérios estatísticos utilizados para avaliar a variância dentro de cada cluster. Apresenta como principal desvantagem a dificuldade em se computar uma partição ótima, em decorrência das inúmeras possibilidades existentes. Em comparação com a família anterior, apresenta como vantagem o fato de utilizar os dados reais, e não uma matriz de similaridade (ALDENDERFER; BLASHFIELD, 1984).

A obtenção do melhor resultado na aplicação destes métodos depende basicamente de quatro etapas: a escolha das variáveis, na qual a forma de seleccionar, a normalização dos dados e a verificação de colinearidade são pontos críticos, do algoritmo adotado, da determinação do número de clusters e da validação do resultado. Escolher as variáveis com base na teoria, e preferencialmente confirmadas por especialistas, e utilizar métodos que permitam ajustes para altas correlações entre as variáveis, que utilizem a distância Mahalanobis D^2 , por exemplo, são procedimentos que permitem aprimorar a escolha das variáveis na primeira etapa. As escolhas do algoritmo e do número de clusters podem ser aprimoradas, pela utilização conjunta dos algoritmos: “Hierárquico por Aglomeração” e “Partição Iterativa”. Os primeiros permitem a obtenção do número de clusters da partição inicial, enquanto que os outros são menos sensíveis a existência de “pontos fora da curva” e aprimoram a homogeneidade dos clusters obtidos. A validação do resultado tem na divisão dos dados de entrada em conjuntos distintos, e conseqüente aplicação da clusterização aos subconjuntos, testes de significância de variáveis externas, teoricamente relacionadas com os clusters, porém não utilizadas em sua obtenção, e a triangulação de métodos, isto é, a aplicação de métodos alternativos, preferencialmente que apresentem fraquezas diferentes, sobre o mesmo conjunto de dados (EVERITT, 1979 ; ALDENDERFER; BLASHFIELD, 1984 ; KETCHEN; SHOOK, 1996).

Finalizando a discussão sobre análise de clusters, cabe ressaltar algumas das propostas de Ketchen e Shook (1996) sobre o uso da análise de clusters em pesquisas de gerenciamento estratégico. A utilização da triangulação através do uso de múltiplas técnicas para garantia de que o resultado obtido seja um artefato metodológico, isto é, a concordância entre dois ou três métodos evidencia que o resultado não é produto de um viés do método. Métodos que permitam ponderar a importância das variáveis de entrada, como as técnicas de clusterização monotônicas, adequando o resultado da pesquisa ao “mundo real”. Tal ponderação, principalmente quando utilizando a opinião de especialistas com vivência, como gerentes e executivos, na indústria sob foco, permitiria a obtenção de resultados

mais realistas. Além disto, a validação dos resultados por parte destes mesmos especialistas permitirá um aprimoramento das técnicas de análise de clusters.

2.2.2. Sistemas Neuro-Fuzzy

O modelo proposto para a identificação de grupos estratégicos é baseado em redes neurais artificiais e sistemas de inferência fuzzy. Estas técnicas computacionais foram propostas no fim da década de 50 e durante a década de 60, constituindo-se em uma abordagem alternativa à inteligência artificial, cujos representantes mais conhecidos são os sistemas especialistas, na construção de sistemas computacionais inteligentes, sendo denominada inteligência computacional (*soft computing*). Passaram a ter maior aplicação a partir de 1974, com a utilização da lógica fuzzy em sistemas de controle, sendo a década 80 caracterizada pelo seu “redescobrimto”, principalmente para as redes neurais artificiais. A sua implementação, de forma a constituir-se em um sistema único, permitindo a combinação da capacidade de aprendizado das redes neurais com o poder de interpretação lingüístico dos sistemas de inferência fuzzy, é denominado sistema neuro-fuzzy (VELLASCO; LOPES, 1999). As arquiteturas adaptativas de redes neurais, e funcionalmente equivalentes aos sistemas de inferência fuzzy, também podem ser consideradas sistemas neuro-fuzzy, ao preservar, através de uma arquitetura adequada, a capacidade de interpretação lingüística característica dos sistemas fuzzy (JANG; SUN; MIZUTANI, 1997). Na primeira parte desta seção são apresentados os conceitos básicos de Redes Neurais, seguida dos conceitos dos Sistemas de Inferência Fuzzy, e finalizando com a apresentação das arquiteturas de sistemas neuro-fuzzy aplicadas a clusterização.

2.2.2.1 Redes Neurais Artificiais

Uma rede neural artificial, ou simplesmente rede neural, pode ser vista como um sistema projetado para modelar, de forma simplificada, a forma como o cérebro humano, utilizando-se de neurônios, as células de processamento, e das conexões sinápticas, realiza as tarefas que lhe são apresentadas. Nesta seção são discutidos seus conceitos gerais, correlacionados ao escopo deste trabalho, com o

objetivo de permitir uma compreensão sistêmica do modelo proposto. A discussão completa dos assuntos apresentados é encontrada na referência HAYKIN, S., **Redes Neurais – Princípios e prática**, 2^a ed., Porto Alegre: Bookman, 2001.

As diversas arquiteturas de redes neurais desenvolvidas até o momento podem ser representadas por um modelo composto de um bloco de entrada de dados, denominado camada de entrada, um bloco constituído pelas camadas escondidas e um bloco de saída de dados. A figura 3 ilustra este modelo. As camadas escondidas e de saída são constituídas pelas unidades básicas de processamento da informação nas redes neurais artificiais, os neurônios, estando estes conectados entre si, para formar as conexões entre as camadas escondidas, entre as camadas de saída e as camadas escondidas, e em algumas raras arquiteturas, dentro de uma mesma camada escondida. A implantação destas conexões, bem como dos neurônios, através de programas de computador, é que permite a solução de diversos problemas de processamento através da simulação, ainda que simplificada, da forma como opera o cérebro.

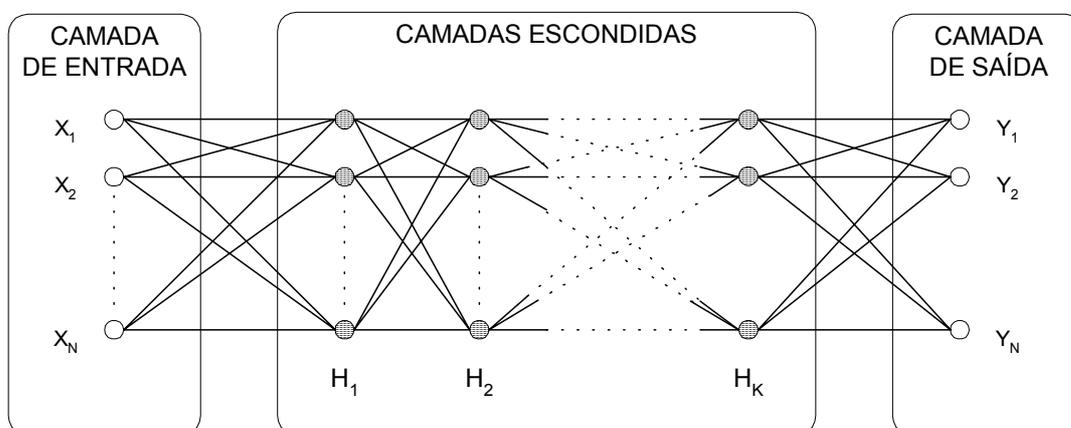


Figura 3 – Modelo básico de arquiteturas de redes neurais artificiais.

O modelo matemático de um neurônio, utilizado para a implantação dos programas de simulação, é ilustrado na figura 4.

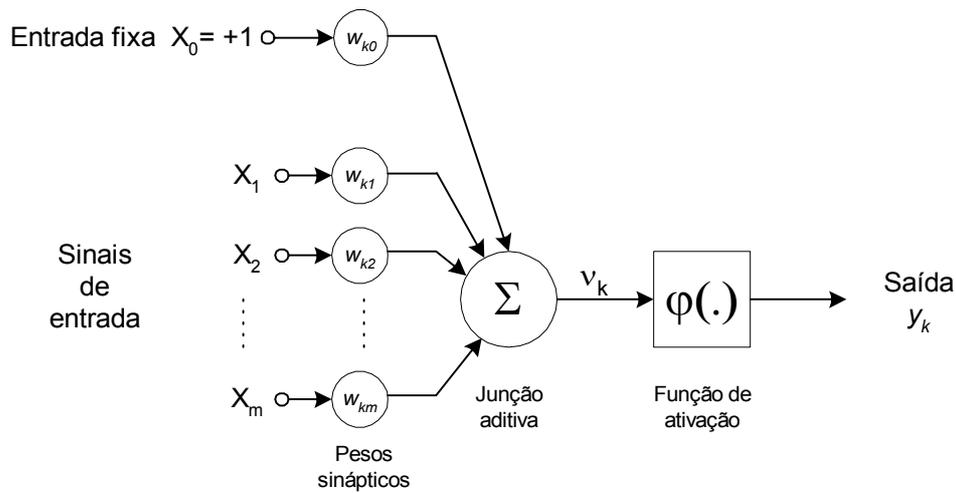


Figura 4 – Modelo matemático não linear de um neurônio.

Os principais elementos deste modelo são o seu conjunto de sinapses, ou elos de conexão, uma junção aditiva, também denominada função de propagação, e a função de ativação. O conjunto de sinapses atua como um dispositivo de entrada para o neurônio, isto é, para cada conexão que é “ativada” no neurônio é atribuído um peso, identificado na figura 4 como w_{ki} e denominado peso sináptico, para ponderar a conexão ativa podendo assumir valores negativos. A junção aditiva serve para efetuar a combinação das conexões ativas, sendo normalmente implementada por um somatório, daí o seu nome. A função de ativação, também denominada função restritiva, uma vez que limita a amplitude da resposta do neurônio, atuando como o dispositivo de saída do neurônio, assumindo normalmente valores entre zero e um. Estes elementos são utilizados na realização de tarefas da seguinte forma: através da camada de entrada, os neurônios da primeira camada escondida são estimulados. A resposta destes neurônios, resultado dos valores de seus pesos sinápticos, da junção aditiva e da função de ativação, ativam os neurônios da próxima camada escondida, e assim sucessivamente, ou diretamente a camada de saída. A camada de saída responde a esta ativação fornecendo a resposta final da rede. Observa-se que o ponto principal deste processo são os valores dos pesos sinápticos de cada conexão, de cada camada. Estes valores são o conhecimento armazenado pela rede e são definidos durante a fase de treinamento, ou de aprendizado, da rede neural.

Por treinamento entende-se a utilização de uma amostra de dados representativas do ambiente de operação da rede para estabelecimento dos valores dos pesos sinápticos, através de métodos computacionais, denominados algoritmos de aprendizado. Uma vez treinada a rede, quando submetida aos estímulos do ambiente, fornecerá as respostas de acordo com o seu conhecimento armazenado. Reside aí a principal propriedade das redes neurais, a sua adaptabilidade. Se o ambiente mudar, as respostas começam a apresentar diferenças em relação ao resultado observado. Para o ajuste da rede, a realização de um novo treinamento, permitirá a adaptação da mesma ao novo ambiente.

O aprendizado das redes neurais acontece através de um processo iterativo de ajustes aplicados a seus pesos sinápticos e níveis de polarização. Idealmente, a rede se torna mais instruída sobre o seu ambiente após cada iteração do processo de aprendizado, sendo este processo realizado pelos algoritmos de aprendizagem. Estes algoritmos podem ser classificados pela forma como é formulado o ajuste de um peso sináptico de um neurônio, a regra de aprendizado, e pela forma como a rede, vista como uma estrutura peculiar de interligações de neurônios, se relaciona com o ambiente. Este último se refere a um modelo do ambiente no qual a rede opera, também conhecido como paradigma de aprendizagem. Quanto à regra de aprendizado, podem ser: por Correção de Erro, Baseada em Memória, de Boltzmann, Hebbiana e Competitiva, destacando-se, para efeito de esclarecimentos quanto ao modelo neuro-fuzzy adotado, esta última. Nesta regra, durante o processo de aprendizado, os pesos sinápticos são ajustados de forma a criar um “ambiente de competição” entre os neurônios de saída, de forma que apenas um será ativado como resposta a dado de entrada. Tal característica torna esta regra adequada para redes neurais que se destinem à classificação de padrões. A figura 5 ilustra a habilidade de realizar a tarefa de agrupamento de uma rede neural que utilize esta regra de aprendizado. Os pontos representam os dados de entrada, enquanto as cruzes os vetores de pesos sinápticos de três neurônios de saída. A esfera (a) indica o estado inicial da rede, enquanto a esfera (b) o estado final.

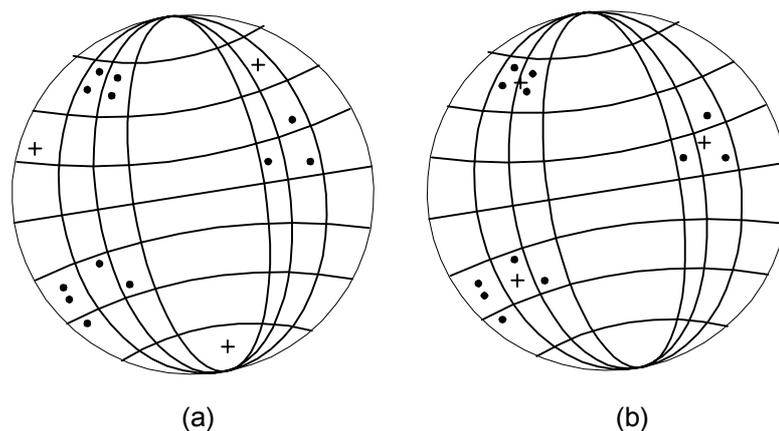


Figura 5 – Visão geométrica da regra de aprendizado competitivo.

A grande restrição desta regra é que os dados de entrada utilizados no treinamento localizem-se em grupamentos razoavelmente distintos. Caso contrário, a rede pode tornar instável mais a um determinado padrão com o mesmo neurônio de saída. Quanto ao relacionamento com o ambiente podem ser classificados em supervisionadas ou não-supervisionadas. No aprendizado supervisionado, o conhecimento sobre o ambiente permite que os ajustes da rede sejam realizados tomando como referência padrões entrada-saída conhecidos. No aprendizado não-supervisionado a rede apresenta recursos de, a partir da regularidade estatística dos dados de entrada, desenvolver a habilidade de formar representações internas para “codificar” as características da entrada. A aprendizagem por reforço e a auto-organizada são os dois tipos de aprendizado não-supervisionados mais utilizados. No primeiro, o aprendizado se dá por tentativa e erro. A cada iteração, o resultado recebe uma avaliação do ambiente sem instruções de correção, apenas se esta certo ou não. Na auto-organizada o processo busca identificar padrões significativos ou características nos dados de entrada, sem a existência de um padrão de referência.

A natureza do presente trabalho, na qual a existência de eventuais grupamentos estratégicos é conhecida a priori, aponta para a utilização de uma arquitetura de rede com aprendizado não-supervisionado. Uma arquitetura que se apresenta apropriada à abordagem deste trabalho é a conhecida como mapas auto-organizáveis (*self organized maps*), proposta por Willshaw e Von der Malsburg (HAYKIN, 2001). Os mapas auto-organizáveis têm como questão principal o fato

de que a ordem global de um sistema pode surgir de interações locais. Este fato é observado no cérebro, no qual muitas interações aleatórias entre neurônios vizinhos de uma rede podem se fundir em estados de ordem global e levar a um comportamento coerente na forma de padrões espaciais ou ritmos temporais.

A organização da rede acontece em dois níveis diferentes: o nível Atividade – no qual certos padrões de atividade são produzidos por uma determinada rede em resposta aos sinais de entrada – e o de Conectividade, no qual as forças de conexão – os pesos sinápticos – da rede são modificados em resposta aos sinais neurais dos padrões de atividade. A integração entre estes dois níveis apresenta os seguintes aspectos: (1) através de realimentação positiva, para se obter a auto-organização ao invés da estabilização, regida pelos princípios de que estas modificações dos pesos sinápticos tendem a se auto-amplificar; (2) existe um ambiente de competição entre as sinapses, decorrentes da limitação de recursos, que resulta na seleção de algumas, que crescem mais vigorosamente que outras; (3) as modificações em pesos sinápticos tendem a cooperar, isto é, a presença de uma sinapse mais vigorosa tende a reforçar o ajuste de outras sinapses, apesar da competição global da rede; (4) a ordem e a estrutura nos padrões de informação representam informação redundante adquirida pela rede na forma de conhecimento.

A arquitetura para implementar estes princípios é formada de tal forma que, a camada de saída é arranjada em uma grade, normalmente unidimensional ou bidimensional, e utiliza a regra de aprendizado competitiva. Assim, os neurônios desta grade “competem” para serem ativados de forma que apenas um neurônio da grade se torne seletivamente sintonizado a um determinado padrão, ou classes de padrões, de entrada no decorrer de um processo de aprendizado. O mapa auto-organizável constitui-se desta forma em um mapa topográfico dos padrões de entrada no qual as localizações espaciais, as coordenadas, dos neurônios da grade são indicativas das características estatísticas intrínsecas contidas nos padrões de entrada. Os dois modelos básicos dos mapas auto-organizáveis são os propostos por Willshaw e Von der Malsburg e por Kohonen, sendo a diferença básica a sua aplicação. O primeiro destina-se a mapeamentos nos quais a dimensão da saída é igual à dimensão da entrada, enquanto que o outro apresenta na “compressão” das

dimensões de entrada, isto é, dimensões menores da saída, uma de suas grandes qualidades. Esta compressão é de grande valia, principalmente quando a dimensão do conjunto de entradas compromete a implementação computacional do modelo ou uma dificuldade na compreensão dos resultados. A figura 6 ilustra um arranjo do modelo de Kohonen com uma camada de saída formada por uma grade bidimensional de cinco neurônios.

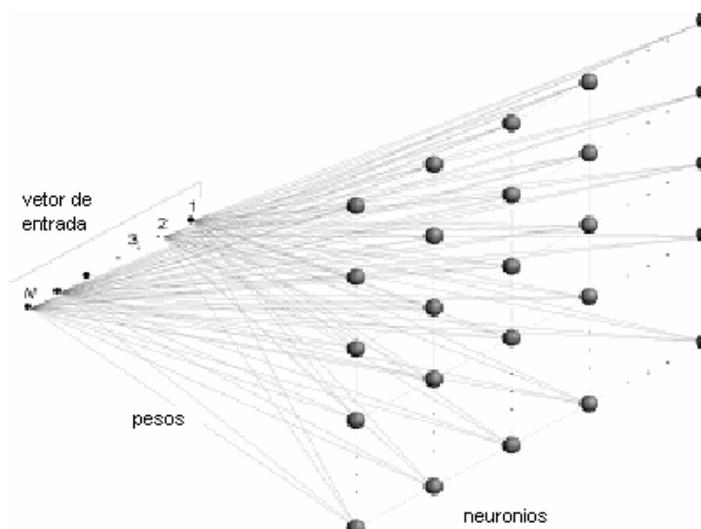


Figura 6 – Estrutura de uma rede neural para realização do modelo de Kohonen.

2.2.2.2 Sistemas Fuzzy

Um sistema fuzzy, ou nebuloso, pode ser caracterizado pela capacidade que um programa de computador, ou um equipamento, de estimar funções e controlar processos e sistemas baseado nas descrições, não raramente parciais, do comportamento dos mesmos. O conhecimento heurístico normalmente é obtido através de especialistas, ou como já acontece, através de redes neurais e outros sistemas fuzzy. Foram desenvolvidos a partir do trabalho de Lotfi Zadeh em 1965 sobre conjuntos nebulosos, no qual foram estabelecidas as bases para a modelagem matemática destes sistemas (KOSKO, 1992). Nesta seção são discutidos seus conceitos gerais da Lógica Fuzzy e dos Sistemas de Inferência, correlacionados ao escopo deste trabalho, com o objetivo de permitir uma compreensão sistêmica do modelo proposto. A discussão completa dos assuntos apresentados é encontrada na

referência JANG, J. – S. R., SUN, C.-T., MIZUTANI, E., **Neuro-Fuzzy and Soft Computing – A Computational Approach to Learning and Machine Intelligence**, New Jersey: Prentice Hall, 1997.

2.2.2.2.1 Conjuntos Nebulosos e a Lógica Fuzzy

A teoria dos conjuntos fuzzy, ou nebulosos, fornece uma base sistemática de cálculo para a criação de regras, as componentes básicas dos Sistemas de Inferência Fuzzy, que podem então ser implementados através de programas de computador. O primeiro conceito é o de conjuntos nebulosos. Na teoria de conjuntos tradicional, a identificação de seus elementos é suficiente para caracterizá-los. Nos conjuntos nebulosos esta caracterização deve ser complementada com um valor para indicar o quanto determinado elemento pertence ao conjunto, sendo este valor definido pela função de pertinência que caracteriza o mesmo. Assim, em um conjunto nebuloso, os elementos são representados por um par ordenado, o elemento e uma medida, esta sempre variando de zero, para não pertencente, até um para completamente pertencente, de quanto o mesmo pertence ao conjunto. Tal característica é que confere aos conjuntos nebulosos a flexibilidade para a modelagem dos sistemas fuzzy, ou seja, uma determinada variável pode pertencer a mais de um conjunto, diferenciados pelos valores determinados pelas funções de pertinência de cada conjunto. Estas funções de pertinência podem ser discretas ou contínuas, devendo apresentar a forma que melhor caracterize o conjunto nebuloso. As formas mais comuns são a triangular, trapezoidal, gaussiana e a função sino. Cabe aqui uma observação para o fato de que se uma determinada variável apresenta uma curva que caracterize sua distribuição de probabilidade, tal informação pode ser utilizada para a construção de um conjunto nebuloso. A figura 7 apresenta uma ilustração das formas de função de pertinência triangular, trapezoidal e gaussiana (JANG; SUN; MIZUTANI, 1997).

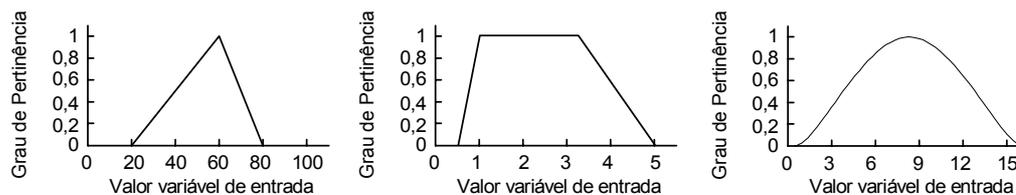


Figura 7 – Formas mais comuns de funções de pertinência.

Outro conceito básico para a construção de sistemas de inferência fuzzy é o de variáveis lingüísticas e, conseqüentemente o de valores lingüísticos. Foi proposto por Zadeh como uma abordagem alternativa à modelagem do pensamento humano para o tratamento de sistemas complexos. Tal abordagem permite, ainda que de forma aproximada, resumir a informação e expressá-la na forma de conjuntos nebulosos e não de conjuntos numéricos. As variáveis lingüísticas são caracterizadas por uma quintupla ordenada composta por: (1) o nome da variável; (2) o conjunto de termos, isto é, o conjunto de valores, ou termos, lingüísticos; (3) o universo de discurso, o intervalo dos possíveis valores numéricos da variável; (4) as regras de sintaxe para a geração dos valores lingüísticos; (5) a regra semântica que associa a cada valor lingüístico uma função de pertinência para cada elemento do conjunto de termos. Assim, por exemplo, a variável lingüística idade pode ser representada por: (idade; {muito nova, nova, meia idade, velha, muito velha}; [0, 100] anos; {0 a 15 anos é muito nova, 13 a 25 anos é nova, 23 a 45 anos é meia idade, 40 a 70 anos é velha, 70 a 100 anos é muito velha}, $\{\mu_{mN}, \mu_N, \mu_{MI}, \mu_V, \mu_{mV}\}$).

São oportunas duas observações na descrição da variável lingüística. A primeira é que μ_{mN} , μ_N , μ_{MI} , μ_V e μ_{mV} são as funções de pertinência associadas a cada um dos termos lingüísticos, podendo apresentar qualquer uma das formas ilustradas na figura 7 ou utilizar algum formato real já previamente conhecido. A outra é que os intervalos estabelecidos para cada termo apresentam interseções. Tais interseções têm sua interpretação dependente das formas das funções de pertinência, porém, apenas como exemplo, pode-se ter que: a idade de 14 anos é muito nova com grau de pertinência 0,05 e nova com grau de pertinência 0,1. É na

definição das variáveis lingüísticas que a experiência dos especialistas contribui para a construção do modelo.

Para completar os conceitos de conjuntos nebulosos, cabe um breve comentário sobre as operações entre os elementos destes conjuntos. Como nos conjuntos numéricos, os conjuntos nebulosos apresentam operações básicas como união, interseção e complemento. A diferença está no fato de que tais operações devem envolver, além dos elementos, a função de pertinência. São definidos então os pares de operadores denominados norma-T e norma-S, ou conorma-T, para representar as classes de operadores que obtêm as funções de pertinência da interseção e da união de dois conjuntos nebulosos. Exemplos para o par, norma-T / norma-S são as operações: mínimo / máximo, e / ou e soma / produto (JANG; SUN; MIZUTANI, 1997).

A Lógica Fuzzy é um sistema de regras que permite lidar com a informação contida em variáveis lingüísticas, pertencentes a um conjunto nebuloso. Por exemplo, o nível de aplicação de um recurso “X” é *alto* ou a forma de gerenciamento deste setor é *altamente* participativo. Três conceitos são básicos na construção deste sistema de regras: a relação fuzzy, a regra se-então fuzzy e o raciocínio fuzzy.

As relações fuzzy são conjuntos fuzzy – n-uplas ordenadas, formadas a partir de n conjuntos de variáveis, e uma função de pertinência – que associam aos elementos de um conjunto n-dimensional um grau de pertinência. São matematicamente representadas através de matrizes de relacionamento. A grande aplicação destas relações fuzzy são as composições que podem ser realizadas entre as várias relações, definidas sobre diferentes conjuntos n-dimensionais, para gerar relações distintas das formadoras. Entre estas composições, destacam-se a Máx-min e Máx-produto. Na primeira, a função de pertinência resultante é composta a partir dos valores máximos de um conjunto, formado a partir dos valores mínimos dos resultados das relações utilizadas na composição. Na outra, os valores máximos são extraídos de um conjunto formado pelo produto dos resultados das relações utilizadas na composição.

As regras se-então, também conhecidas como implicações fuzzy, são expressões utilizadas para estabelecer a relação entre variáveis lingüísticas. Por exemplo, a expressão “se a velocidade está alta e se o sinal está vermelho então pise fundo o freio do veículo” correlaciona as variáveis lingüísticas: velocidade, condição do sinal e acionamento do freio. Os trechos “se a velocidade está alta” e “se o sinal está vermelho” são denominados antecedentes, ou premissas, enquanto que o trecho “então pise fundo o freio do veículo” é chamado de conseqüente, ou conclusão. Estas expressões são matematicamente modeladas através das relações fuzzy e suas composições.

Uma visão importante do conjunto de regras se-então diz respeito ao espaço de entrada. Os antecedentes das regras criam uma partição do espaço de entrada estabelecendo uma região para cada variável lingüística. Os conseqüentes por sua vez, descrevem o comportamento dos sistemas de inferência nestas regiões, através de seus constituintes, funções de pertinência ou equações lineares das entradas, por exemplo, que dependem do tipo de sistema de inferência escolhido. O fator importante da partição do espaço de entrada diz respeito à relação entre as quantidades de variáveis de entradas, de funções de pertinência e de regras. O número de regras cresce exponencialmente em função das quantidades de funções de pertinência e de variáveis de entrada, podendo inviabilizar a implantação do sistema de inferência fuzzy. Este problema, denominado curso dimensional, é tratado através da adoção de estratégias diferentes de partição do espaço de entrada. Os tipos mais comuns de partição do espaço de entrada são: (1) a partição grade, a mais simples, porém que gera o maior número de regras; (2) em árvore, que permite especificar as regiões de forma específica, porém com prejuízos ao entendimento e ao número de funções de pertinência; (3) a difusa, no qual o espaço de entrada é dividido em regiões em que se tem a maior possibilidade de ocorrências das variáveis de entrada. Neste tipo normalmente é necessário o conhecimento da característica entrada / saída desejada ou esperada. Existem outros tipos de partição, mas que, de uma forma geral, são variantes destes três tipos. A figura 8 ilustra estes três tipos básicos, sendo que de (a) a (c) utiliza como

constituintes funções de pertinência. De (d) a (f) os constituintes são combinações lineares das variáveis de entrada (JANG; SUN; MIZUTANI, 1997).

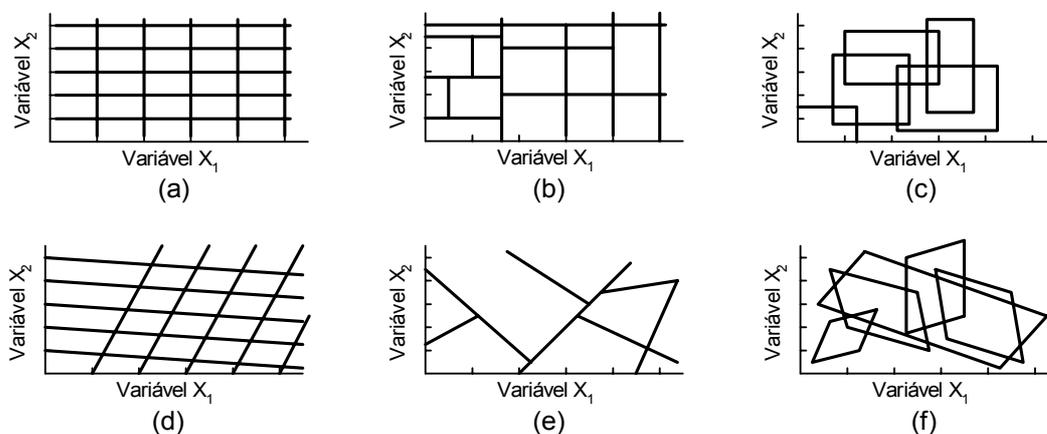


Figura 8 – Exemplos para os tipos de partição de um sistema com duas variáveis.

O raciocínio fuzzy, também denominado raciocínio de aproximação, é um procedimento de inferência para a obtenção de conclusões a partir de um conjunto de regras se-então e fatos conhecidos. O conceito base do raciocínio fuzzy é o da regra de composição de inferência, utilizada na modelagem matemática do raciocínio fuzzy. Por esta regra, se dois conjuntos numéricos estão correlacionados por uma função, o conjunto fuzzy associado ao conjunto numérico de saída é igual à composição entre a relação fuzzy associada ao conjunto numérico de entrada e a relação fuzzy associada à função que correlaciona os dois conjuntos. A regra básica de inferência da lógica tradicional, é que a verdade do conseqüente é inferida a partir da verdade do antecedente. Esta regra, denominada *modus ponens*, é aplicada ao raciocínio fuzzy para obter o procedimento de inferência fuzzy denominado *modus ponens* generalizado. Por este procedimento, com base em fatos e um conjunto de regras se-então se pode inferir uma conclusão. Por exemplo: utilizando a regra “se a velocidade está alta e se o sinal está vermelho então pise fundo o freio do veículo” e o fato de que a velocidade está mais ou menos alta tem-se como conclusão “pisar mais ou menos fundo o freio do

veículo”. Este raciocínio é modelado matematicamente através da composição de relações fuzzy para permitir a sua implementação através de programas de computador (JANG; SUN; MIZUTANI, 1997).

2.2.2.2 Sistemas de Inferência Fuzzy

Os conceitos estabelecidos pela teoria de conjuntos nebulosos e da lógica fuzzy são utilizados na geração de programas de computador aplicados à solução de problemas como controle e automação, classificação de dados, análise de decisões, sistemas especialistas e robótica entre outras. Tais programas são conhecidos como sistemas de inferência fuzzy. Os principais elementos de um sistema de inferência são: (1) a base de regras, que contém o conjunto de regras se-então; (2) um dicionário de funções de pertinência utilizadas pelas regras; (3) o procedimento de inferência, também denominado agregador; (4) caso a saída necessite ser numérica, um processo para conversão de fuzzy para número, denominado “defuzzificação”. A figura 9 ilustra um diagrama básico de um sistema de inferência fuzzy.

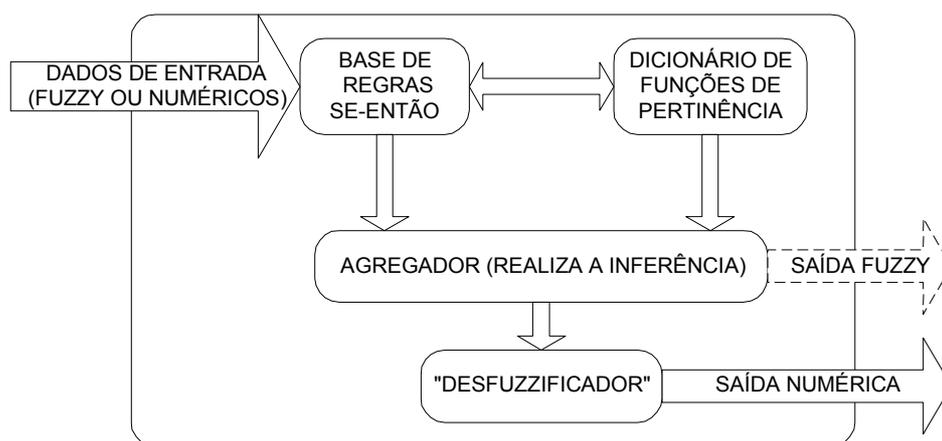


Figura 9 – Diagrama básico de um sistema de inferência fuzzy.

Apenas para efeito descritivo, se pode dividir os sistemas de inferência fuzzy em dois grandes grupos: aqueles que utilizam apenas o raciocínio fuzzy e aqueles que são baseados em arquiteturas neuro-fuzzy. Os sistemas baseados na

arquitetura neuro-fuzzy pertencem ao grupo conhecido como ANFIS (*Adaptive Neuro-Fuzzy Inference Systems*) e tem na capacidade de armazenar o conhecimento a sua principal vantagem, apresentando também grande aplicação na identificação de regras. A grande maioria dos sistemas ANFIS se baseia no aprendizado supervisionado, isto é, necessitam de um padrão já conhecido, não sendo adequada a sua utilização neste trabalho. Do grupo que utiliza apenas o raciocínio fuzzy destacam-se os modelos de Sugeno, também conhecido como modelo TSK, e o de Mandani. O primeiro foi proposto para desenvolver uma abordagem sistemática na geração de regras fuzzy, com base em um conjunto de pares entrada / saída conhecidos. Neste modelo o consequente das regras é uma função, normalmente uma constante, ordem zero, ou uma combinação linear, de primeira ordem, dos valores numéricos da entrada, não sendo necessário o uso do processo de “desfuzzificação” da saída. A necessidade de dados conhecidos da saída neste modelo inviabiliza, pelo menos em um primeiro momento, a sua aplicação neste trabalho. O modelo de Mandani se baseia em um conjunto de regras estabelecido pela experiência de especialistas. O procedimento de inferência utiliza normalmente como norma-T e norma-S os operadores mínimo e máximo (JANG; SUN; MIZUTANI, 1997).

A figura 9 mostra que o resultado da aplicação do modelo de Mandani gera como saída um conjunto fuzzy. Para convertê-la em um valor numérico deve ser aplicado um processo de “desfuzzificação”, dos quais os mais utilizados são: centro de área, o mais utilizado, bisetor de área, média de máximos, menor de máximos e maior de máximos, embora existam outros métodos mais flexíveis. Estes processos não são submetidos a uma análise matemática rigorosa, sendo a maioria dos estudos baseados em resultados experimentais (JANG; SUN; MIZUTANI, 1997).

2.2.3. Sistemas Neuro-Fuzzy aplicados a Identificação de Clusters

Embora os Sistemas de Inferência Fuzzy permitam a representação do conhecimento através de suas regras, sua capacidade em lidar com as mudanças do ambiente é limitada, uma vez que pequenas mudanças no conjunto de entradas

podem levar a alterações no resultado de um sistema de inferência fuzzy. A incorporação da capacidade de aprendizado das redes neurais aos Sistemas de Inferência Fuzzy, resultando em um modelo neuro-fuzzy, permite a superação destes limites. Nesta seção serão discutidos brevemente o algoritmo de clusterização fuzzy FCM e a utilização de redes neurais que utilizam a arquitetura dos mapas auto-organizáveis, modelo de Kohonen, associada aos conceitos dos conjuntos fuzzy para a identificação de clusters.

O algoritmo de clusterização FCM é uma versão fuzzy do algoritmo K-Means, também denominado C-Means. Neste algoritmo, são associados aos dados “pertencentes” a um determinado cluster graus de pertinência, ao invés de simplesmente uma indicação de pertinência, normalmente 1, ou não, que é o resultado da versão não fuzzy. Com esta característica, um determinado ponto pode estar associado a mais de um cluster. Pode ser resumido nos seguintes passos: (1) arbitrar os graus de pertinência iniciais dos dados; (2) calcular as coordenadas iniciais dos centros utilizando os dados e seus respectivos graus de pertinência; (3) computar o valor de uma função de custo, encerrando caso seja satisfeito o critério de parada, previamente estabelecido; (4) caso contrário, realiza do cálculo de novos graus de pertinência e retoma a iteração, o passo (2). O algoritmo FCM também não apresenta garantia de convergência para uma solução final ótima, sendo a performance dependente da escolha inicial dos graus de pertinência (JANG; SUN; MIZUTANI, 1997). Uma forma para superar esta limitação, está na utilização de redes neurais de aprendizado não-supervisionado, como o modelo auto-organizável de Kohonen. A característica desta arquitetura, na qual alguns neurônios da camada de saída são ativados de acordo com os dados de entrada, de forma não-supervisionada, permite utilizá-la na análise de clusters. Entretanto, o modelo de Kohonen apresenta, além dos problemas descritos para o FCM, problemas quanto à estratégia de aprendizado. Bezdek, Tsao e Pal propuseram um modelo integrando o modelo de Kohonen e o algoritmo FCM, utilizando os graus de pertinência calculados por este na taxa de aprendizado do modelo de Kohonen. O modelo resultante foi testado com o um conjunto de dados bem conhecido e comparado com os resultados dos algoritmos C-Means, FCM e de Kohonen, apresentando melhoras quanto à convergência assim como erros na identificação

dos clusters. Este modelo foi denominado Fuzzy Kohonen Clustering Networks ou FKCN (BEZDEK; TSAO; PAL, 1992). Outros modelos também foram propostos por (CHI; KUO; TENG, 2000), para a realização de segmentação de mercado de clientes de cartão de crédito, e (MITRA ; PAL, 1994), testado no reconhecimento de sinais de voz, no caso vogais do idioma hindu, utilizando o modelo de Kohonen e os conceitos de conjuntos nebulosos.